Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Telecommunication Engineering

Efficient Control, Routing, and Wavelength Assignment in Loss-Less Optical Burst Switching Networks

Disertation thesis

Ing. Miloš Kozák

Ph.D. programme: Electrical Engineering and Information Technology
Branch of study: Telecommunication Engineering
Supervisor: Doc. Ing. Leoš Boháč, Ph.D.

Prague, March 2015

Thesis Supervisor:
Doc. Ing. Leoš Boháč, Ph.D.
Department of Telecommunication Engineering
Faculty of Electrical Engineering
Czech Technical University in Prague
Technická 2
160 00 Prague 6
Czech Republic

Copyright © March 2015 Ing. Miloš Kozák

Declaration

I hereby declare I have written this doctoral thesis independently and quoted all the sources of information used in accordance with methodological instructions on ethical principles for writing an academic thesis. Moreover, I state that this thesis has neither been submitted nor accepted for any other degree.

In Prague, March 2015

.....
Ing. Miloš Kozák

Abstract

The thesis deals with the problem of loss-less Optical Burst Switching (OBS) paradigm viability for future all-optical networks. This problem is decomposed into two problems: a) data transmission b) network control, and each problem is individually verified.

The data transmission problem is related to the statistical multiplex on the optical domain hence the initial work is on preparation of a tool for simulation of loss-less OBS network, and Car-OBS (CAROBS) architecture is selected in this thesis. The CAROBS architecture is unique because allows buffering of contenting bursts in electronic domain. This buffering, however, degrades network performance parameters and increases capital expenditure (CAPEX) and operating expense (OPEX). In order to find the relation between performance parameters and node load, an elementary node behaviour was studied using the CAROBS models implemented with OMNeT++ simulator. This experiment shows relation between node load and requirement of opto-electronic (O/E) blocks which cause increase of CAPEX and OPEX. Subsequently, the experiment shows that it is possible to define a routing strategy in CAROBS network in such a way that no O/E blocks are necessary. This observation led to the formulation of an optimization algorithm which provides routing strategy such that minimum number of O/E blocks are used in a given network. The linear programming was selected because it can provide an optimal solution, but the formulation is not scalable due to the problem complexity.

Heuristics were not used because a rigorous method is necessary in order to formulate a claim on the loss-less OBS viability. Therefore, a special case of this complex formulation was selected. In this special case flows are not merging, i.e., Stream-line effect (SLE), and based on this case an optimization algorithm called SLE-RWA was formulated and implemented. This routing strategy is then verified using OMNeT++ models, and it was showed that loss-less OBS can achieve considerably higher wavelength efficiency than what is possible with Optical Circuit Switching (OCS) networks. Therefore, this routing strategy outperforms OCS systems.

Then, the focus was shifted towards to the OBS network control. Network control does not depend on loss-less properties, but the control proposed in this thesis is based on CAROBS and software-defined networking (SDN), i.e., Software-defined CAROBS (SD-CAROBS) is defined. This network control concept uses logically centralized controller and redefined OpenFlow (OF) messages in order to deliver necessary information for CAROBS framework control. The SDN based control is vital because it allows unifying optical infrastructure and service control in an Internet service provider (ISP) network. Therefore, using SD-CAROBS control mechanism allows to control OBS network in a seamless way that is still not possible in OCS networks.

Keywords: Optical burst switching, loss-less, operational research, column generation, simulations.

Acknowledgements

First and foremost, I would like to express my most sincere and earnest gratitude to Dr. Brigitte Jaumard. Without her guidance, support, and enthusiasm throughout my research which led to the end of my Ph.D. studies. Her immense knowledge, helpful suggestions and comments, as well as encouragements, made this research possible. I am grateful to her for the valuable time she put into all the discussions we ever had, either in the lab or overseas through Skype.

Very special acknowledgement goes to my family. You should know that your support and encouragement was worth more than I can express on paper.

I would like to express my gratitude to my supervisor Dr. Leoš Boháč for his never-ending support, and his special way to motivate me towards the next and even deeper research. I remember being grumpy many times after our discussions, but his arguments implying my weak readiness made me think deeper and in advance what can be the next step, rather than just look at the problem and make the first assumption to be the solution.

An extensive time I spent in labs on two continents where I met a whole bunch of brilliant researchers. I think I stole a bit of wisdom from each of them, so thank you very much guys. I mean you Tomáši, Zbyňku, Petrové, Lukáši, Radku, Hoa, Minh, Anh, Tian, Michel, Mahdi, Hong Hai, and of course Florian!

Personally, I would like to thank the most awesome girls in the world: Christina, Jennifer, Barbie, and Sara for making every single day the best day while back in Montréal.

At last, I would like to thank grants: SGS10/275/OHK3/3T/13, SGS11/124/OHK3/2T/13, SGS12/186/OHK3/3T/13, SGS13/200/OHK3/3T/13, and TA02011015 because they significantly helped in finishing of my research.

Un tas de pierres cesse d'être un tas de pierres, dès qu'un seul homme le contemple avec, en lui, l'image d'une cathédrale.

–

A rock pile ceases to be a rock pile the moment a single man contemplates it, bearing within him the image of a cathedral.

Antoine de Saint-Exupéry

List of Tables

List of Figures

List of Acronyms

ACC access delay. 75, 79, 89, 90

AON all-optical networking. 97, 115, 116, 118

AP aggregation pool. 45, 72, 73, 101–103

API Application Programming Interface. xi, 14, 95, 98, 99

AQ aggregation queue. xi, 44, 45, 72, 73, 79, 101–103

ASON Automatically Switched Optical Network. 3, 11

AWG arrayed waveguide grating. 10

BBP burst blocking probability. 48, 49, 53, 75, 79, 89–91, 111

BD buffering delay. ix, 55–58, 60, 75, 79, 89–91, 94, 111

BHP burst header packet. 8, 9, 11, 28–35

BLP burst loss probability. 51, 71

BP buffering probability. ix, 48, 49, 53, 55–61

CAPEX capital expenditure. iv, 47, 60, 62, 65, 70, 71, 93, 97, 112, 115–117

CAROBS Car-OBS. iv, ix, xi, 42–45, 47–54, 57–62, 65, 67, 68, 70–75, 79, 80, 89, 90, 92–94, 96, 99–101, 103, 110, 113, 115–118

CG column generation. 84, 93, 94, 115, 116

CHP CAROBS header packet. x, xi, 43–45, 72–74, 76, 97, 99, 100, 103

CPU Central Processing Unit. 116

CSMA-CA Carrier Sense Multiple Access with Collision Avoidance. 42

DPID DataPath IDentifier. xi, 100–103

DWDM dense wavelength division multiplex. 6

E/O electro-optical convertor. ix, xi, 50, 100

E2E end to end delay. 79, 89, 90, 94, 110, 111

EOT extra offset time. 44

FDL fiber delay line. 7, 8, 38–40, 42, 43, 48

FIFO first-in first-out. 8

FPGA Field Programmable Gate Array. 32

FSM finite state machine. 14

GMPLS Generalized Multi-Protocol Label Switching. 3, 11, 12, 95, 96, 105

GRWA grooming routing and wavelength assignment. 50, 51, 59, 62, 71, 115

HPC high performance computing. 15

ILP integer linear program. 65, 66, 68–70, 80–82, 84, 85, 107, 109, 115, 117

IP Internet Protocol. 3, 5, 8, 32, 99, 105

IQS Input QS. ix, 49, 50

ISP Internet service provider. iv, 1, 4, 89, 95, 98, 100, 113, 116

JET just-enough-time. 34, 36, 37, 45, 53, 89, 109, 110

JIT just-in-time. 34, 35, 109

LAN Local Area Network. 3, 4

LAUC Latest available unused channel. 36, 37

LAUC-VF Latest available unused channel with void-filling. ix, 36, 50, 53, 89, 110

LP linear program. 14, 17, 20, 23, 47, 71, 80, 107, 109, 115

LTE Long Term Evolution. 1

MAC media access control. 43, 100–102

MAN Metropolitan Area Network. 4, 42

MEMS micro-electro-mechanical systems. 10

MP master problem. 84, 107, 109

MVA mean value analysis. 60, 116

MX Switching matrix. x, 30–34, 38, 41, 43, 45, 67, 68

O/E opto-electronic. iv, ix–xi, 5, 32, 43, 47, 49, 50, 52–54, 58–65, 67–70, 90, 92–94, 100, 111–113, 116

OBS Optical Burst Switching. iv, viii, 5, 8–14, 24, 26, 28–30, 32–34, 37, 39–42, 44, 47, 48, 51, 52, 65, 71, 95, 96, 107, 109, 111, 115–118

OBTN Optical Burst Transport Network. 41

OCS Optical Circuit Switching. iv, vii, 5–12, 28, 42, 47, 48, 50, 90, 93, 96, 105, 111–114, 117

OF OpenFlow. iv, 95–97, 99–102

ONF Open Networking Foundation. 96

OPEX operating expense. iv, 47, 60, 62, 65, 70, 71, 93, 97, 98, 112, 115–117

OPS Optical Packet Switching. 5, 7–12, 32

OPST Optical Packet Switching and Transport. 42, 43

OQS Output Queueing systems. 49, 50

OR operational research. 80, 107

OSPF Open Shortest Path First. 105

OT offset time. 9, 28, 29, 32, 34, 35, 43–45

OTH Optical Transport Hierarchy. 2, 11, 12

OXC Optical Cross-connect. viii, 3, 6–10, 88

PDH Plesiochronous digital hierarchy. 2

PLZT lead lanthanum zirconate titanate. 10

POADM Packet Optical Add/Drop Multiplexing. 40, 41

PP pricing problem. 86, 88, 107

QoS Quality of Service. 9, 31, 96

QS Queueing systems. ix, xiii, 49, 50, 52

RAM random access memory. 5, 8, 9, 38

RMP restricted master problem. 84–86, 107, 109

ROADM Reconfigurable Optical Add Drop Multiplexer. 3

RWA routing and wavelength assignment. vii, xi, 12, 41, 44, 47, 50, 51, 66, 67, 95, 97–99, 101, 102, 104–106, 109–114, 117, 118

SAP Service Access Point. xi, 98, 99, 104, 105

SD-CAROBS Software-defined CAROBS. iv, xi, 96–106, 109–114, 116–118

SDH Synchronous Digital Hierarchy. 2, 11, 12

SDN software-defined networking. iv, 95–101, 104, 105, 113, 116, 117

SECR secondary contention ratio. 75, 79

SLE Stream-line effect. iv, 41, 47–49, 52, 53, 57, 61, 71–75, 79, 80, 82, 83, 85, 86, 88, 91, 93, 115, 117

SOA semiconductor optical amplifier. 10

TCP Transmission Control Protocol. 99

TWC tuneable wavelength converter. 10

UDA universal data analyzer. 26, 119, 122

WAN Wide Area Network. 4, 6

WC wavelength conversion. 4, 6, 71

WDM wavelength division multiplex. ix, 2–5, 28, 30, 54, 58–62, 64–66, 114, 115

WR-OBS Wavelength-Routed Optical Burst Switching. 41, 42

WWW World Wide Web. 1

Contents

Chapter 1

Introduction

In the early 1980s a revolution in telecommunications networks began and spawned by the use of a relatively unassuming technology: fiber optic cable. Since then, the tremendous cost savings and increased network quality has led to many new services delivered over networks using optical fibers, and many new benefits of very fast network are only beginning to be realized. In the early 90", telecommunication networks were used for voice calls, dial-up was the greatest development, and first very low rate international peerings were emerging in order to enable data communication between states. Today, on the other hand, one can use an ordinary cellphone in order to access Internet with download speed around 150 Mbps[1], and recent standards specify even higher data rates. This download speed is almost negligible compared to optical access networks where the transmissions speeds can be symmetric and download speed can be 10 Gbps [1].

These extraordinary speeds in access networks are necessary due to many factors: the tremendous growth and penetration of the global communication network called Internet and its service World Wide Web (WWW) which is the main source of information for Internet users. However, the WWW is starting to be overcome by applications using communication capabilities provided by Internet. The impact of applications on the required bandwidth is caused by their number. Each user can have more devices where some applications can be installed (phone, tablet, etc.), and there can be more than just one application connected to the Internet at the same time. Therefore, the amount of data transmitted in ISP's networks is continuously increasing even though it was assumed the bandwidth usage is going to culminate because the Internet penetration has been saturated in developed world [2].

Thanks to the technological advances, cost of the bandwidth has been continuously reducing. Additionally, the deregulation of the classical market of voice and data services initiated competition which in turn has resulted in lower end user costs and faster deploy-

[1]This is current maximal download speed in a Long Term Evolution (LTE) network.

ment of new technologies and services. This deregulation has also resulted in creating a number of new start-up service providers as well as companies providing equipment to these service providers. As one can see, this development drives the lust for more bandwidth in the network, so the bandwidth requirement has no sign of abating in the near future.

These factors have driven the development of high-capacity optical networks and their remarkably rapid transition from research laboratories into commercial deployment. The twisted pairs would never allow transmitting as high data rates as one wavelength can support in optical networks for long distances. Dealing with such a huge capacity optical networks provide an efficient mechanism must be defined. Therefore, this thesis is dealing with optimal traffic routing, and control mechanisms in optical networks of future.

1.1 Optical networks

The explosive extension of networks and services as we know today would not be possible, unless optical fiber attenuation was reduced by Mr. Kao in 60". This achievement brought to Mr. Kao nobel price in 2009, and massive improvement of all types of communications for mankind. Optical fibers allowed to increase the distance between exchanges and speed of all communication channels. Originally, point-to-point connections were constructed as simple replacement of coopered pairs. When manageability of such an approach was saturated, new optical network systems started to be designed. First, it was Plesiochronous digital hierarchy (PDH) and later Synchronous Digital Hierarchy (SDH) systems. Synchronous transmission mechanism, e.g., SDH and Optical Transport Hierarchy (OTH), can almost endlessly scale in terms of transported bandwidth due to the containers multiplexing. This scaling, however, is conditioned by extra capacity necessary for channel control. This extra capacity pushes the channel capacity requirements on quality of transmitted signal very close to the physical limitations where optical signal experiences propagation problems hence the scalability is very difficult which means increase of production capacity is very expensive. Additionally, synchronous networks are based on ring topologies, so connection to backbone ring networks through access rings. Customers are demanding more services, better latencies, and transfer more data. In order to meet all the transmission requirements, a new paradigm was needed. Nowadays, optical networks are constructed using synchronous technologies, *Ethernet* protocol on link layer, and WDM on physical layer. Recent studies mention all-optical networking as future of optical networks. The development of optical networks can be split into three generations [3]:

- **1st generation** of optical networks simply replaced copper wires with optical fibers.

The 1st generation optical networks started to use optical fibers in order to increase link capacity, and in point-to-point installations. However, optical fibers are different compared to copper wires. Optical fibers allow a number of optical signals at different carrier wavelengths to be simultaneously carried by the same fiber.

- **2nd generation** of optical networks take into account differences between fiber and copper, e.g., fibers allow transmission of multiple wavelengths – WDM. Another very important aspect of this generation is an optical control plane. Optical control plane is focused on protocols for unified network configuration. The result is Generalized Multi-Protocol Label Switching (GMPLS), Automatically Switched Optical Network (ASON). Both these control concepts are used in order to drive the Reconfigurable Optical Add Drop Multiplexer (ROADM), OXC. Thanks to the automatic configuration of intermediate elements a long term optical connection across an optical network can be established. In other words, capacity of a wavelength is the minimal granularity in this generation of optical networks.

- **3rd generation** of optical networks avoids electronic switching, redefines control, and uses more mature OXC in order to allow often switching. This generation of optical networks handles traffic very same way as packets are handled in Internet Protocol (IP) networks that means no permanent circuits are reserved for long term connections and grooming is carried out optically. This assumption has implications in modification of network control. There is no waiting for optical circuit establishment, but all intermediate switching nodes are configured just in time. This means every wavelength is not devoted to certain communication, but each wavelength is shared by some flows in time, i.e., 3rd generation of optical networks introduce subwavelength granularity routing.

Optical networks vary in the size, which can be quantified geographically and by the amount of nodes. Each size of an optical network has implications into the robustness of used technology. In order to classify network size and make selection of network nodes easier, optical network can be classified as:

- **Local Area Network (LAN)** connecting network devices over a relatively short distance. A networked office building, school, or home usually contains a single LAN, and occasionally a LAN will span a group of nearby buildings. Ethernet as link protocol is used in majority of LANs. LAN contains electronic switches that are interconnected by optical fibers, but all switching is carried out in the electronic domain. Most often used topology is star or hierarchical campus.

- **Metropolitan Area Network (MAN)** usually connects two or more LANs in a single network that has geographic area larger than single LAN, but smaller than the region. MANs are mostly built in cities or towns to provide a high speed data connection and usually owned by a single large organization. A MAN bridges a number of LANs with a fiber-optical links that act as a backbone.

- **Wide Area Network (WAN)** is a computer network covering larger geographical area such as a state, province or country. It provides a solution to companies or organizations operating from distant geographical locations who want to communicate with each other for sharing and managing central data or for general communication. WAN network interconnects MAN and LAN networks such that they can communicate with each other, and be geographically far away from each other. Generally, WAN is managed by one body who is referred to ISP. WANs are mainly optical networks because the bandwidth that must be supported, 100 Gbps and more, is not possible to transmit using other technologies. Nowadays, WANs are designed with technologies of first and second generation of optical networks, but the technologies of the third generation of optical networks are highly desired due to power savings.

In all generations of optical networks, the power consumption is a significant concern [4]–[9] because the amount of transmitted bandwidth is very high in WAN, and the power consumption increases proportionally to the transmitted bandwidth [10], [11]. The power consumption, economic crisis, and requirement for low latencies accelerated deployment of WDM systems and research of third generation of optical networks.

1.2 All-optical switched networks

The second and third generation of optical network uses a technique called *bypass* [12] which enables to switch an incoming light at certain wavelength to defined output port, and if wavelength conversion (WC) is used wavelength can be changed as well. Optical networks using optical bypass can considerably reduce propagation delay due to electronic processing is avoided. Efficient use of optical network is achieved due to WDM. The ramification of all-optical network is captured in Figure 1.1 where one can see that the interoperability with legacy or non-optical technologies is realized by edge nodes that convert communication protocols of the adjacent network to the communication system of the all-optical network.

The only drawback compared to the first generation of optical network that used ring topologies is mesh topology. The protection and restoration is more complex in networks

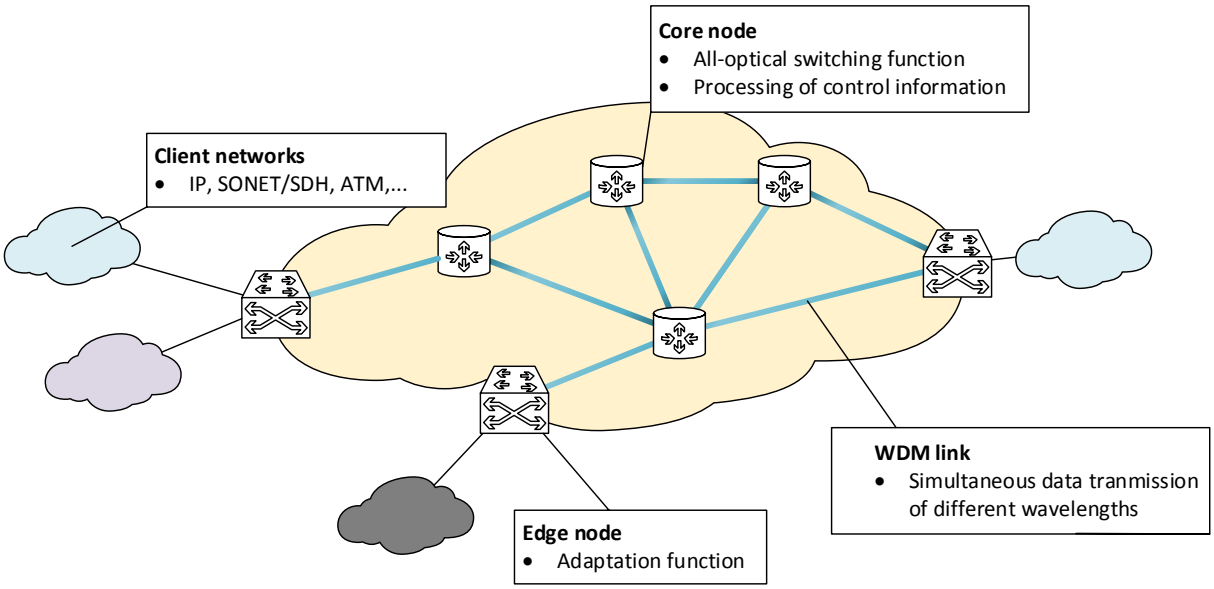


Figure 1.1: Generic architecture of optical DWDM network.

using mesh topologies compared to networks with ring topologies. On the other hand the same optical fibers can be used; consequently, upgrade from the first generation of optical network to the all-optical network can be evolutionary. Also, due to the WDM only some wavelengths can be used for mesh topologies, i.e., virtual topologies. Over the time, whole network infrastructure can convert from ring virtual topologies to mesh virtual topologies.

There are three popular switching paradigms that have been proposed for the all-optical networks: OCS, Optical Packet Switching (OPS), and OBS [13]. In OCS an end-to-end *lightpath* is established between a source node and a destination node for the entire session which can last a few seconds and more. Lightpaths avoid O/E conversions; therefore, the end-to-end delays considerably decreases and grooming is not possible. In the OPS, packets from the client network are directly switched in the optical domain along the path as it is common in IP networks, but the IP networks use electronic switches with electronic memory for congestion avoidance. The optical random access memory (RAM) memories are not ready for production deployments hence packet loss can occur in OPS networks. The third switching paradigm is OBS proposed in order to overcome the deficiencies of both OCS and OPS. Also, OBS is seen as a technology to support all-optical switching in the Internet before OPS is ready. The benefit of OBS and OPS over OCS is no need to dedicate the capacity of a wavelength for the entire session between a pair of nodes. This is extremely important for sessions that do not fully use capacity of a wavelength. Moreover, OBS is more viable than OPS because packets, for OBS bursts, does not need to be buffered at the intermediate nodes, so OBS can deliver lower

latencies. In the following section, the all-optical paradigms are discussed in more detail to give reader better overview of transmission principles.

1.3 Optical Circuit Switching

The principle of OCS network is a connection between two nodes created for the duration of the transmission between source and destination node. This connection is also called a *lightpath* because the connection uses one or more wavelengths [14].

Lightpaths are created by reserving dedicated wavelength on every link along physical path through network. Lightpaths can be either static or dynamic. Some systems of second generation of optical networks set up lightpaths manually, and often last for several days. Other systems configure lightpaths automatically according to the traffic requests. The lightpath switching is carried out at by OXC at every intermediate node. OXC is responsible for the all-optical switching of the data that are carried on an incoming wavelength (usually denoted by λ) in a given input port, to an outgoing wavelength in an output port. However, a OXC has no sense what is transported in the switched wavelength. If the OXC is equipped with WC then the lightpath may be converted from one wavelength to another along the route. Otherwise, the *wavelength continuity* constraint is applied.

The OCS technique is well-known from the wavelength routed networks. Wavelength routed networks use wavelength as an identifier used for routing or switching. OCS is a mature technology, already deployed in the WAN. Although OCS can be considered as significant improvement over the first generation of optical networks it has some disadvantages. The main disadvantage is *large granularity*. The smallest switching entity in OCS networks is the wavelength, but since one wavelength of dense wavelength division multiplex (DWDM) network can support more than 100 Gbps then it is complicated to use this wavelength efficiently. Typically, the OCS connection lasts a few seconds and more. The connection setup and release is in milliseconds [15]. OCS is good for static lightpath connections with traffic requests known in advance. *Bursty traffic* can cause problems because it can leave some wavelength capacity unused or is buffered at the edge when the traffic is higher than expected. Therefore, the OCS can be very inefficient when traffic dynamically changes. Most importantly, the *number of connections* in a network is usually much greater than the number of wavelengths, and the required bandwidth for a connection is much smaller than the capacity of a wavelength. Therefore, without electronic grooming it is neither possible nor efficient to allocate one wavelength to every connection [13], [16]. This results in so called optical clusters connected by nodes that do electronic grooming.

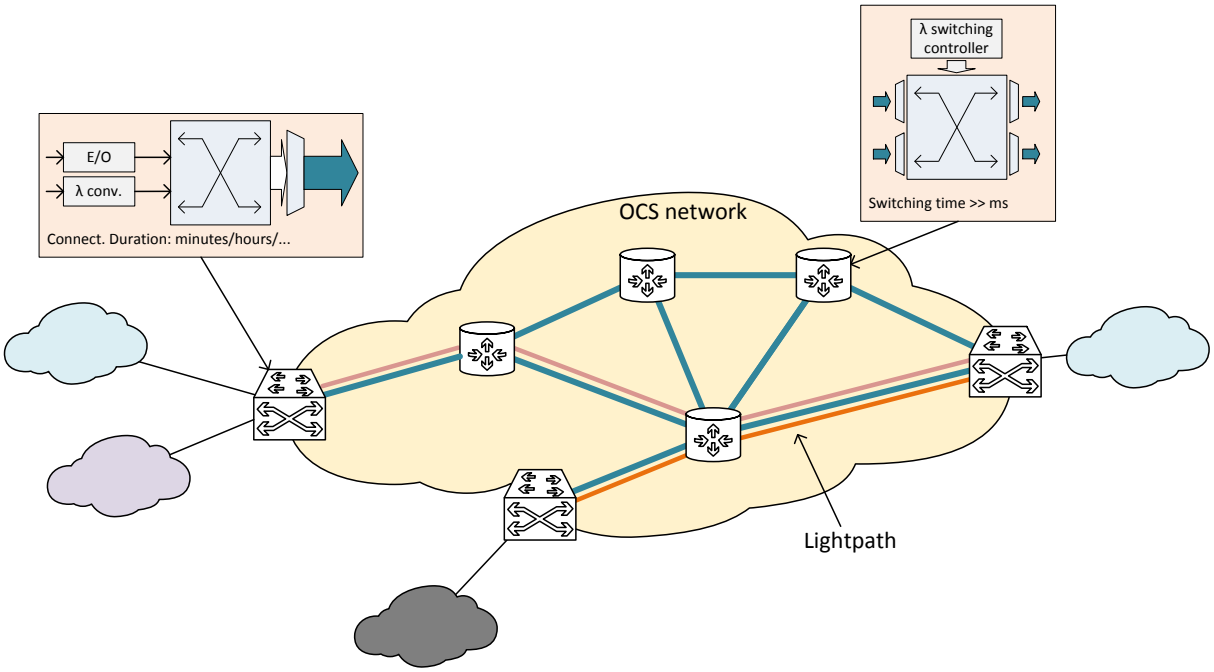


Figure 1.2: OCS network (Picture adapted from [17]).

Typical function of OCS network is captured in Figure 1.2 where you can see that lightpaths sharing one link must be put onto different wavelength.

1.4 Optical Packet Switching

The principle of OPS network is optical transmission of packets that contain only one packet or frame from the client network [18]. Packets are switched and routed independently throughout the network without any prior reservation. This should be done entirely in the optical domain. However, due to technological limitations, the research community has shifted the focus from photonic control, which would need quantum computers, to electronic control. When a OPS packet is received in an optical switch, its header is extracted and converted to electrical form. The header is then processed in an electronic controller while the payload is optically delayed using a fiber delay line (FDL), and the whole packet is forwarded after proper switching. OPS brings some technological challenges to current technology such as: *Header processing* which requires the header to be optically extracted from each packet then converted to its electrical form for the electronic processing. This operation must be applied to every single packet, so very fast processor is necessary. *Synchronization* is another challenge in photonic packet switching networks because the contention can be minimized only for equally long packets at switch input [18]. These packets are small, as they contain packet or frames from client network hence the switching of OXC is very often and as a sequel OPS requires very *fast switching*

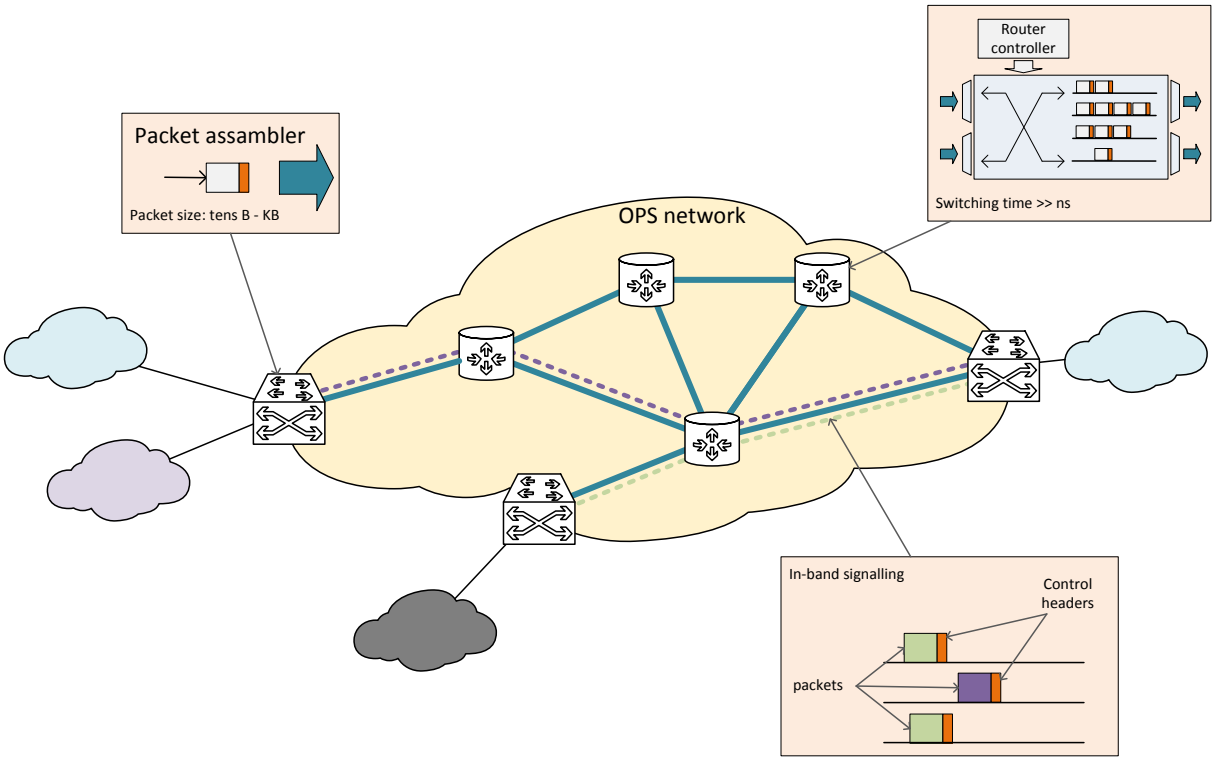


Figure 1.3: OPS network (Picture adapted from [17]).

matrix. OPS does not reserved a lightpath in advance as OCS, so packets may experience contention. Contention in IP networks is typically handled through buffering. However, there is no optical equivalent of RAM and buffering is realized by FDL. The limitation of FDL is that even for small delay relatively long fiber is necessary, FDL works as first-in first-out (FIFO) system with predetermined delay, but FDL degrades the optical signal physical properties such as magnitude and delay.

The objective of OPS is to enable packet switching at the same rate as OCS ensures, but with considerable higher link efficiency. This would bring an impressive increase in network utilization by means of statistical multiplexing for bandwidth sharing. Typical bandwidth sharing is captured in Figure 1.3.

1.5 Optical Burst Switching

OBS is mentioned in literature as a balance between OCS and OPS because it is more feasible than OPS, but provides statistical multiplexing in optical domain, i.e., optical grooming. The principle of OBS is the transmission of the data bursts, which are an transport unit that carries multiple frames or packets from client network, and just in time channel signaling. These bursts are switched all-optically by an OXC using information carried in burst header packet (BHP). The BHP is transmitted on a dedicated control

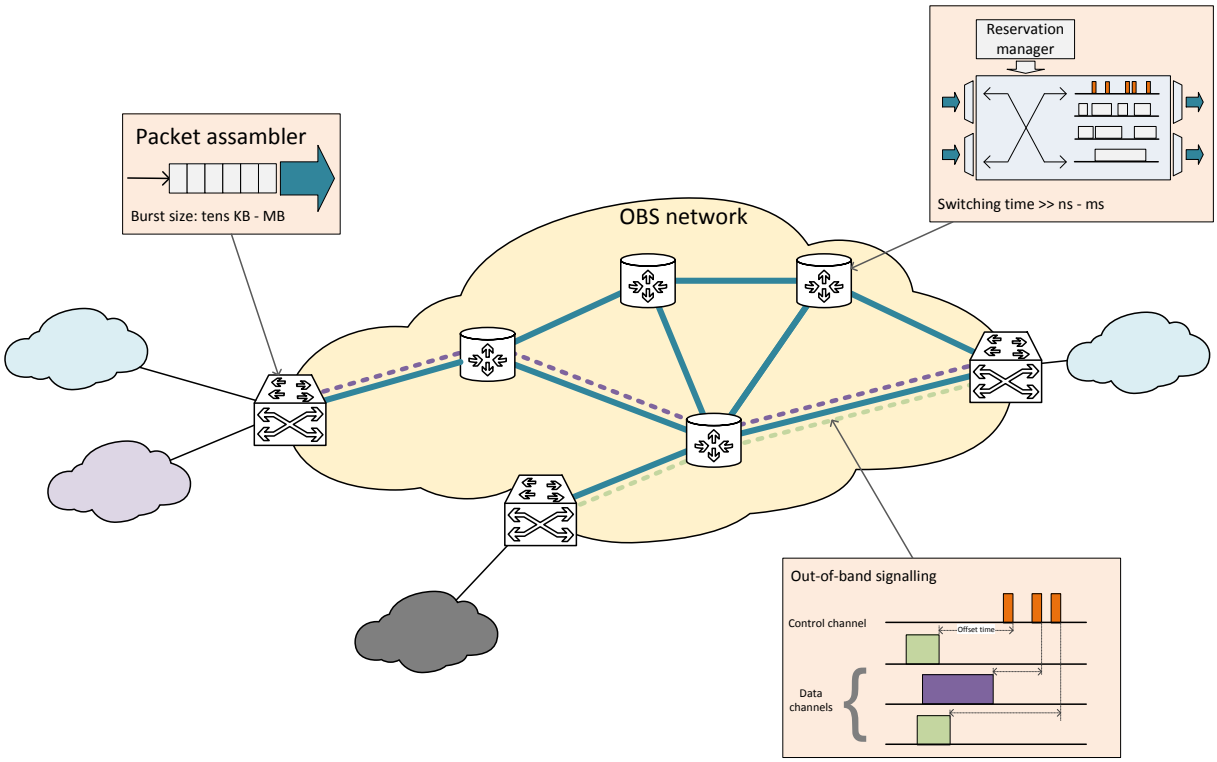


Figure 1.4: OBS network (Picture adapted from [17]).

channel and precedes the burst in order to give some time the core node to process this BHP and configure OXC. Based on the information carried by the BHP the intermediate nodes reserve particular switching resource for the duration of the data burst. This time between BHP and burst is called *offset time (OT)*, and it is proportional to the number of hops and Quality of Service (QoS). An OBS network is depicted in Figure 1.4.

OBS allows dynamical reservation of paths, called *flow-paths*, for small and very specific periods of time, optical grooming, using wavelength as a shared medium, and efficient bandwidth allocation. The efficient bandwidth allocation was unachievable by OCS technology. Due to bursts are larger than OPS packets, processing overhead is reduced and the technology requirements for switching fabric and processing are not so hard. The duration of a typical burst can last from some μs to several hundreds of ms. OBS presents enormous advantages over OCS, and is easier to implement compared with OPS. Nevertheless, OBS is not ready for deployment, and if there is an technological jump enabling optical RAM memories OBS will never go to market.

1.6 Differences among all-optical paradigms

Each paradigm has some benefits and limitations, so it is not possible to give credit to only one representative. In the following text, the main differences will be discussed.

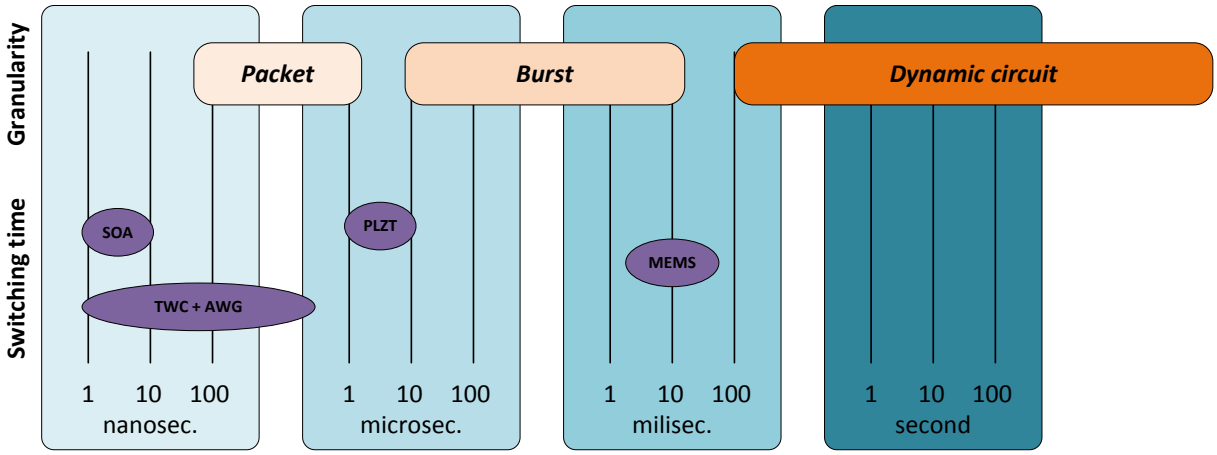


Figure 1.5: Comparison of transport unit duration versus technology necessary in order to realize OXC for the appropriate paradigm.

1.6.1 Switching fabrics

The aforementioned paradigms differ in the required switching time, and different technologies to implement OXC can be used. The switching time has direct implication to the selection of switching technology, see Figure 1.5. However, the prices of switching technologies are not same hence the price of switching fabric varies based on the used system.

The switching function in OCS nodes can be achieved with commercially available optical switch technologies such as micro-electro-mechanical systems (MEMS) [19]. The relatively long switching times of MEMS are not an obstacle for low-dynamic OCS operation. On the other hand, a dynamic character of OBS and OPS requires a fast-switching operation. The ns switching times are achievable with commercially available arrayed waveguide grating (AWG) and tuneable wavelength converter (TWC) [20] technologies or the semiconductor optical amplifier (SOA)[21] technology. The lead lanthanum zirconate titanate (PLZT) technology [22] with μs switching times can be a good choice for OBS, and the SOA can be very good option for OPS where there is a massive amount of quasi-simultaneously arriving OPS packets, so switching must be very fast in order to minimize packet loss due to blocking.

1.6.2 Adaptability and scalability

Adaptability addresses ability to adapt client flows to different bit-rates and data formats in order to transport them through the all-optical network. In the OCS architecture such adaptation is achieved either by using an E/O conversion or with wavelength conversion for optical tributary signal. However, OCS only works with granularity of whole wave-

length. This means even analog signal can be transmitted trough an OCS networks. On the other hand the OBS and OPS are more complex and are focused on data communication; therefore, it is not possible to transmit an analog continuous signal through OBS or OPS network without its packetization at an edge node.

OBS contains assembly/aggregation function which is responsible for burst formation at the ingress edge node and for the burst disassembly at the egress node. The design issue is the proper choice of a burst aggregation scheme with its associated parameters, namely the size of the burst and acceptable burst formation time which relates to the optical grooming scheme. On the other hand OCS does not allow any optical grooming at all, so granularity relates to the wavelength capacity. This disadvantage is partly reduced by synchronous systems, e.t. SDH and OTH, but those systems can do grooming only in the electronic domain, so the optical network is not transparent, but translucent [23].

Paradigms for networks of future must respect the network efficiency. OBS and OPS achieve higher wavelength efficiency OCS

1.6.3 Control plane

OCS has been used in networks for a while, so there are control protocols which are part of GMPLS/ASON architecture (connection management, protection/restoration, etc.). GMPLS/ASON is also mentioned for management of OBS and OPS network in literature [24].

GMPLS was designed for OCS networks originally. In order to create a lightpath, each node along the path must acknowledge required resources then the lightpath is set usually for long time duration. OPS benefits from the GMPLS architecture because allows creation of accurate routing tables. OPS packet contains both data and header which is used for routing. This structure of a OPS packet allows to route packets separately.

On the other hand in the OBS, there is the offset time between BHP and burst which limits routing decisions. Basically, any change of routing of already sent burst is limited which means alternative path can be selected only if the alternative path contains equal or less number of hops. Moreover, every burst duration is very short hence GMPLS acknowledging scheme would be a bottleneck, so the GMPLS is not very good candidate for OBS network control.

At the same time, the control of OBS is not standardized, and every research team tackles it uniquely, and sometimes very vaguely.

1.6.4 Routing

Routing deals with traffic distribution over the network. The dynamic traffic distribution is very difficult task for OCS network because the quasi-static network state translates to the RWA problem [25]. The traffic variation impacts the network lightpath configuration, and this is very demanding for control protocol, e.g., GMPLS, to distribute routing information to the network. Some changes can cause excessive waiting of some traffic at edge node, but can never cause frame loss inside of the all-optical network due to contentions. The traffic distributions is well studied for OCS networks.

The OBS and OPS can tackle the dynamic changes better thanks to the optical grooming. Generally, these paradigms can use network more efficiently for the same offered load. However, when the traffic dynamically changes both paradigms deal with this situation differently. When the network load increases the burst or packet contention occurs more often. Contention is a situation when two or more incoming burst or packets intend use the same egress port. OPS deals with this situation using rerouting mechanism using another shortest-path towards to the destination. OBS is in more difficult situation because rerouting is limited, see section 1.6.3 hence contention can cause burst loss, retransmission, and less efficiency eventually. The routing in OBS network is well studied [26].

Network survivability of OCS networks is well studied [27]; however, this problem is still not enough addressed in OBS and OPS. Also, this is one of the reasons why OBS networks are not widely deployed.

1.6.5 Conclusion

This section compared the three all-optical paradigms. None of these paradigms is a leader. OCS is mature technology that is currently deployed, but OCS is usually combined with electronic systems, e.g., SDH, OTH, Ethernet, in order to provide grooming and GMPLS in order to signal end-to-end channel through this protocol stack.

OBS and OPS paradigms are still in research stage. OPS is in difficult situation because relies on technologies which are still not available for production deployment. The routing problem is well studied for OBS, but control is not defined, and GMPLS is not suitable.

Overall, OBS is still the closest candidate for deployment because it combines benefits of OBS and OCS. On the other hand, OBS is not trusted by research community due to high burst loss. Nevertheless, there are field trials and commercial products based on OBS paradigm [28]–[30].

Chapter 2

Aims of the doctoral thesis

This dissertation thesis deals with the limits of loss-less OBS networks. There are two related areas that need to be verified. It is the *i*) transmission *ii*) control of a loss-less OBS network. Specifically, the targets of this dissertation thesis are:

1. Analysis of loss-less OBS traffic properties, and definition of the working conditions under which the selected loss-less OBS framework can be used. Discussion the impact of loss-less OBS traffic properties onto the loss-less OBS viability.

2. Address the viability of loss-less OBS architecture and optimizations of loss-less OBS networks.

3. Formulate the optimization mechanism of the selected loss-less OBS architecture and show the maximal performance of it.

4. Define a control mechanism which is usable for OBS architecture and compatible with current technology. The proposed control mechanism must easily integrate into currently maintained networks. Also, this control mechanism must allow delivering the maximal network performance.

5. Implement simulations model using an event driven simulator in order to have a tool for verification of studies of the loss-less OBS networks.

Chapter 3

Methods of modelling and evaluation

The goals of this thesis are fulfilled with sequential use of methods according to the Kolb's experimental circle [31]. In this work combination of simulations and optimizations are used, and these methods create a framework allowing to create conclusions about routing in OBS networks, i.e., network flows. Network flows and systems are two very similar terms that influence each other. Systems are usually nodes in a network which support transmission of data in form of flows. Every flow has certain properties that can be quantified using mathematical models. These properties come from the behaviour of edge nodes that follows a protocol of the communication system. The full understating of both systems and flows is conditioned by their deep knowledge, but understanding of network behaviour is even more complex. Therefore, optimizations in the field of networks and flows is very demanding and must be always verified.

In the following sections, and this thesis itself, I am going to focus on combination of simulations and linear program (LP) [32], [33] in OBS networks.

3.1 Simulations

Simulations are an ideal tool in order to increase knowledge of studied system. Simulations represent a real system, but in the computer, so the system can be put into situations that are not even possible in real world.

3.1.1 OMNeT++

OMNeT++ is an object-oriented modular discrete event network simulation framework, and it was used in order to create simulator of the problem studied in this thesis. OMNeT++ was selected because it is OpenSource and provides very good API allowing to create very scalable event driven simulations. Event driven approach was selected because communication protocols are a finite state machine (FSM) that changes its states in time,

i.e., definition of event driven simulations. There are more simulators available, but only OMNeT++ has low entry barrier, is free for academic use, and can be used with high performance computing (HPC) clusters which is very important in order to achieve high accuracy.

OMNeT++ itself is a very generic simulator, and it provides infrastructure tools for writing simulators. One of the fundamental principles is a *component* architecture for simulation models. Models are assembled from reusable components termed modules. Well-written modules are truly reusable, and can be combined in various ways like LEGO blocks. The best approach for writing a module is following UNIX philosophy which emphasizes building short, simple, clear, modular, and extensible code that can be easily maintained and reused.

Modules can be connected with each other via gates, and combined to form compound modules. The depth of module nesting is not limited. Modules communicate through message passing, where messages may carry arbitrary data structures. Modules can pass messages along predefined paths via gates and connections, or directly to their destination; the latter is useful for wireless simulations, for example. Modules at the lowest level of the module hierarchy are called simple modules, and they encapsulate model behaviour. Simple modules are programmed in C++, and make use of the simulation library – inherit simple module object.

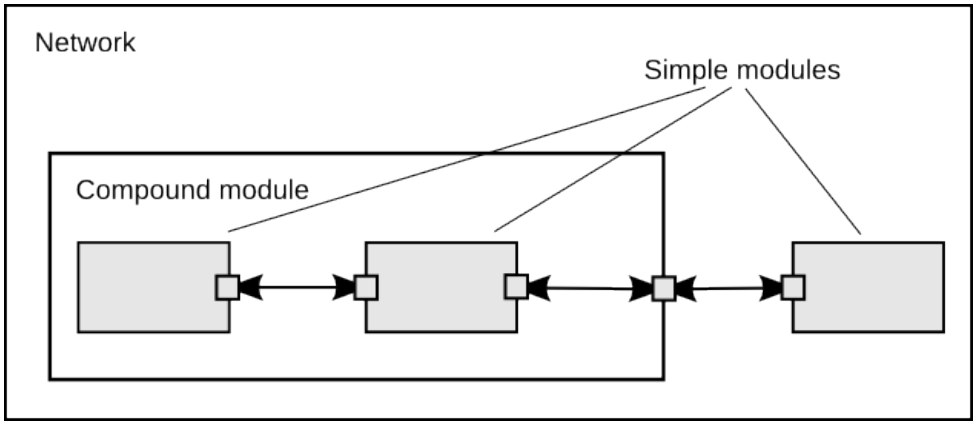


Figure 3.1: Simple and compound module structure. This approach allows creation of complex modules, and keep the simulation models flexible [34].

OMNeT++ simulations can be run under various user interfaces. Graphical, animating user interfaces are highly useful for demonstration and debugging purposes, and command-line user interfaces are best for batch execution at HPC clusters. The simulator is tested on the most common operating systems (Linux, Mac OS/X, Windows), and they can be compiled out of the box or after trivial modifications on most Unix-like operating systems. This is very important because most HPC clusters run on UNIX systems.

OMNeT++ also supports parallel distributed simulation. OMNeT++ can use several mechanisms for communication between partitions of a parallel distributed simulation, for example MPI or named pipes. In general, models do not need special instrumentation to be run in parallel, but some rules need to be followed otherwise it is not possible to configure simulations to be run in parallel [34].

3.2 Mathematical modelling

Mathematical modelling aims to describe the different aspects of the real world, their interaction, and their dynamics through mathematics. It constitutes the third pillar of science and engineering along theoretical analysis and experimentation. One of the reasons for growing success of mathematical modelling is the extension of scientific computation. Scientific computation allows the translation of a mathematical model into algorithms that can be treated and solved by ever more powerful computers. New disciplines started to use mathematical modelling, e.g., information and communication technology, bio-engineering, financial engineering, and so on. As a matter of fact, mathematical models offer new possibilities to manage the increasing complexity of technology. Eventually, the mathematical models can be combined with mathematical optimizations in order to find optimal state of the modelled situation.

3.2.1 Optimizations

Mathematical optimizations are methods allowing finding minimal or maximal value of given function, and they are often used in situations when analytical formulation does not lead to the solution due to high complexity. Additionally, these formulations can be further restricted which makes the analytical solution even more demanding. Majority of engineering problems is possible to solve using at least one optimization technique. Each problem is unique, but may be quantified by certain pattern addressed by a optimization technique. There are three main types of optimization techniques.

Deterministic algorithms are based on analytical mathematical methods. These algorithms require some conditions to be satisfied in order to work efficiently. Problem should be: linear, convex, limited, continuous, analytically described, or objective function contains up to only one extreme point. Deterministic methods give only one solution. Some deterministic optimizations techniques: *Greedy, Hill-climbing, Branch & bound, Depth-first, Broad-first, Best-first, Calculus based.*

Stochastic methods rely on problem randomness. Simply speaking, it is only random approach of finding parameters of objective function, but with the intentions of

finding maximal or minimal value of objective function. Stochastic methods return only one value, but this value is the maximal or minimal from the set of found solutions. The main problem of stochastic methods is they are relatively slow for big problems. It is recommended to use stochastic methods for problems of small degree or for rough guess. Some stochastic methods: *Random search-walk, Simulated annealing, Monte Carlo, Tabu search, Evolutionary computation, Stochastic hill-climbing* .

Mixed methods represent a smart mixture of Deterministic and Stochastic methods which cooperate in order to achieve good results. These methods are very efficient for large-scale problems. Among these methods can be classify: *Matematical programming, Ant colony optimization, Immune system methods, Memetic algorithms, Scatter search & path rel., Particle swarm, Genetic algorithms, Differential algorithms, Soma.*

In the following text I am going to focus on accurate techniques dealing with large-scale optimizations, so the main scope is going to be Mathematical programming. Network flows can be characterised by a set of linear equations, so the main focus is going be Linear programming technique.

Basics of linear programming

A LP is a mathematical formulation of a problem. Problem is defined as a set of decision variables which fully describe the decisions that need to made. Decision variables are also used to define an objective function which is minimized or maximized, and a set of constraints which restrict the decision options. In a linear program, the variables must be continuous, and the objective function and constraints must be linear expressions. An expression is linear if it can be expressed in the form $c_1x_1 + c_2x_2 + \cdots + c_nx_n$ for some constants $c_1, c_2, \ldots, c_n$. For example, $2x_1 + 7x_2$ is a linear expression; x^2 and $sinx$ are not.

Any LP can be seen as a structure composed of three parts *i)* the objective function – (3.1) *ii)* the constraints – (3.2) *iii)* the sign restrictions of the decision variables – (3.3). Sometimes the formulation is obvious, and other times constructing the model is more challenging. An example of LP can be formulated as follows:
Objective function:

$$\min -0.4x_1 + 3.2x_2 \tag{3.1}$$

Subject to:

$$\left.\begin{aligned} x_1 + x_2 &\leq 7 \\ -x_1 - 2x_2 &\leq 4 \\ -x_1 - x_2 &\leq 5 \\ x_1 &\leq 5 \end{aligned}\right\} \quad \text{Constraints} \tag{3.2}$$

$$\left.\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}\right\} \quad \text{Sign restrictions} \tag{3.3}$$

Finding the optimal value of objective function (3.1) is very easy. This set of constraints is based on two decision variables, and using simple geometry one can visualise [1] these constraints. Each constraints represent one function of two variables, so it is easy to draw it. In terms of geometrical visualisation, the sign restriction is a special type of constraint which tackles only one decision variable, but still can be drawn as a line which is perpendicular to the axis representing the variable. The geometrical visualisation of constraints and sign restrictions are depicted in Figure 3.2

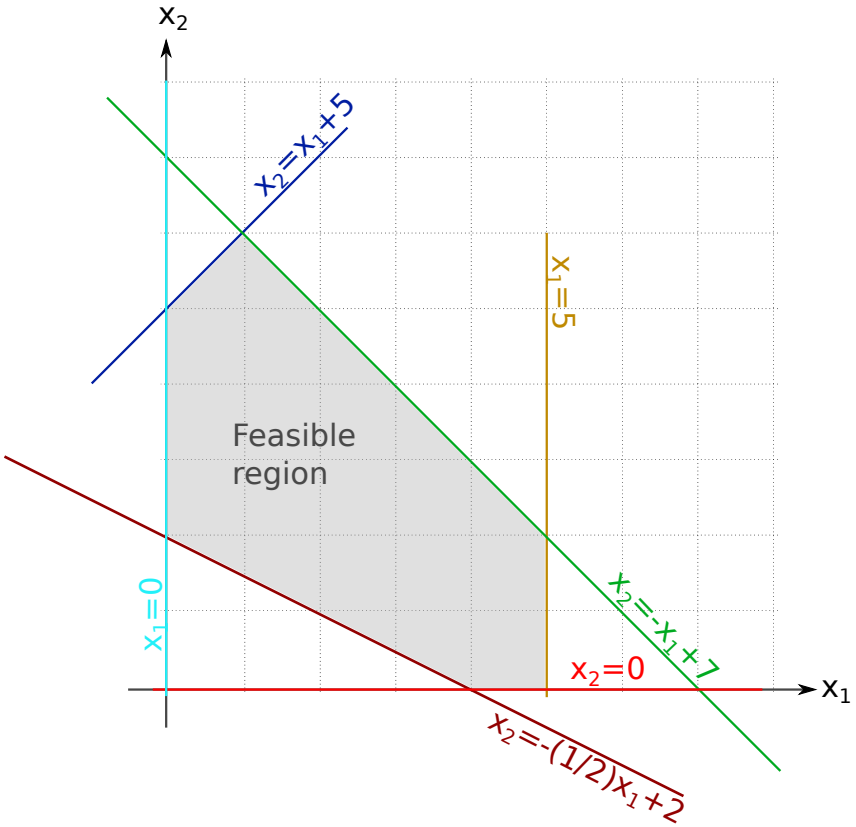


Figure 3.2: Geometrical representation of constraints (3.1) - (3.3).

[1] Finding the optimal value in an visualisation of more than three variables is difficult; however, geometry method still works.

When all constraints are graphed, there is the region where all the constraints are satisfied. This region is called the feasible region. Any point in this region satisfies all the constraints, and is called *feasible*. Any point outside of this region violates one or more of them, and is called *infeasible*.

If a line representing the objective function, see Figure 3.3, is added then the feasible region is split into two areas where the objective function is smaller or bigger. Depending on the optimization approach[2], one side of the feasible region is selected for the further investigation. The objective function (3.1) is an minimization function so the are of interest is on the right-hand side region because variable x_1 is multiplied by a negative coefficient, so the variable x_1 must be maximized whereas the variable x_2 must be minimized.

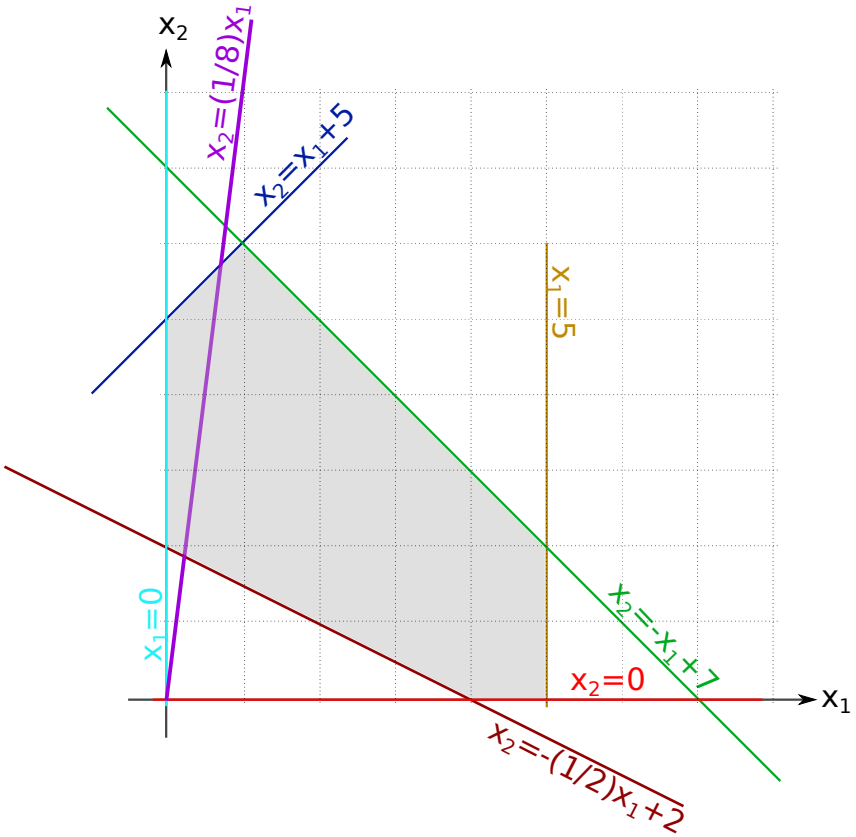


Figure 3.3: Geometrical representation of constraints (3.1) - (3.3) and objective function (3.1).

The maximal distance that the objective function to be pushed toward the boundaries on the right side of the feasible regions represents the minimal value of objective function (3.1). The minimal value of objective function (3.1) is $z^* = -2$ for $x_1 = 5$ and $x_2 = 0$.

[2]Whether the objective function is minimization or maximization

Simplex method

The example in the previous section gave good example how optimization works, and how optimization problem can be described using mathematical formulation. The graphical solution of LP gets on its complexity with the increasing number of variables. In case of three variables lines representing constraints change to planes, and feasibility region is still defined by them, but it is more complicated structure. These feasibility regions are also called polyhedron, see Figure 3.4.

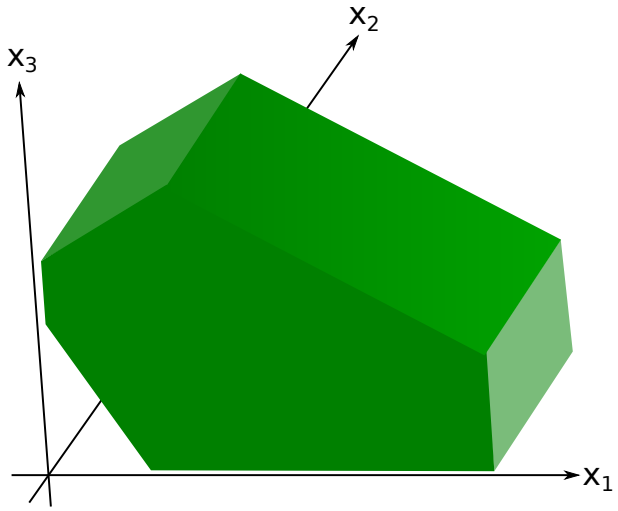


Figure 3.4: Polyhedron representing set of constraints based on thee decision variables.

As in the previous section, the objective function determines border between maximal and minimal values, but the objective function is represented by a plane in the situation. However, the approach of finding the optimal value is the same.

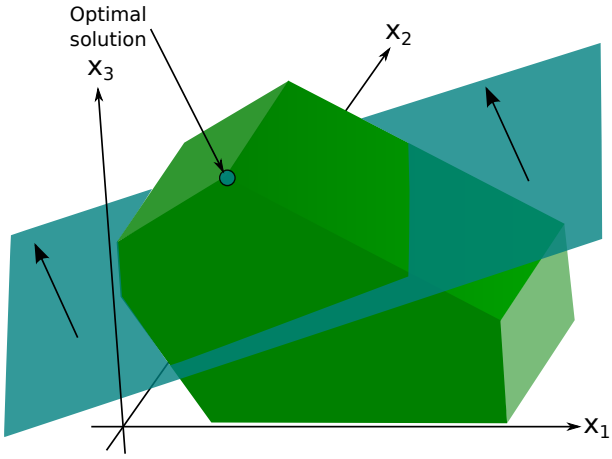


Figure 3.5: Polyhedron representing set of constraints based on thee decision variables, and blue plane representing objective function and direction towards to the optimal value.

The only problem of geometrical method is the number of variables that can be visualised, and analysed by a human. Other problem of this method is the human interaction

hence more advance method was necessary to be developed because the increasing number of constraints increases the number of potentially optimal values, i.e., corners given by constraints.

Using the Gaussian elimination, these corners can be efficiently found, but the problem is their exact position in respect to the objective function plane. Therefore the *Simplex method* was developed. The Simplex method is an efficient method for solving linear programs, and is currently used in most commercial solvers. The simplex method intelligently moves from corner to corner until it can be proven that the solver has found the optimal solution. Each corner that it visits is an improvement over the previous one. Once it can't find a better corner, it knows that it has the optimal solution.

Deeper description of Simplex method is out of the scope of this thesis, and the main goal in this thesis is not to describe mathematical tool, but to use it in network planning. Readers interested into internal function of this mathematical tool refers [32], [35]. Nevertheless, for the sake of the further text the notation for simplex must be mentioned. The simplex method uses matrix representation for variables, so the objective function (3.1), constraints (3.2) , and sign restrictions (3.3) must be rewritten into the standard notation:

$$\min \quad c^T x \tag{3.4}$$

$$\text{subject to :} \quad Ax = -b \tag{3.5}$$

$$x \geq 0 \tag{3.6}$$

3.2.2 Optimizations of network flows

Network flows optimization is very comprehensive discipline because it requires knowledge of Graph theory which is very important for description of the underlying network, Mathematical modelling which is necessary in order to describe network and flows, and Mathematical programming which is used for the optimization in the modelled situation.

Many linear programming problems can be viewed as a problem of minimizing the "transportation" cost of moving materials through a network to meet demands for material at various locations given sources of material at other locations. Such problems are called network flow problems. They form the most important special class of linear programming problems. Transportation, electric, and communication networks provide obvious examples of application areas. Less obvious, but just as important, are applications in facility location, resource management, financial planning, and others.

Two following sections are focused on *minimum-cost network* problem which is the main formulation of flows in terms of linear programming, and *maximal flow problem*

which is the foundation of optimizations carried out in this thesis.

Minimum-cost network

The Minimum-cost network problem requires a graph defining underlying information for decision variables hence it is important to define $\mathcal{V}$ as a set of nodes and a set of links $\mathcal{L}$ which are defined as pairing of nodes such that $\mathcal{L} \subset \{(i,j) : i,j \in \mathcal{V}, i \neq j\}$. Network graph G is then defined as pairing function $G(\mathcal{V},\mathcal{L})$. In the rest of the text all graphs are going to be considered to be directional even if it is not explicitly mentioned.

To specify a network flow problem, b_i is defined and denotes amount of bandwidth being supplied to the network at node $i \in \mathcal{V}$ (if $b_i > 0$ then node is a source, and sink otherwise) for each $i \in \mathcal{V}$. Then the problem is to move the material that sits at the supply nodes over to the demand nodes. The movements must be along the arcs $\mathcal{L}$ of the network $G(\mathcal{V},\mathcal{L})$. Transported material can not leave or enter the network except at supply and demand nodes respectively. This is called *flow conservation*, *nodal balance*, or *Kirchhoff equations*:

$$\sum_{i \in \mathcal{V}} b_i = 0 \tag{3.7}$$

In order to find the minimum cost path, along which material should move, parameters c_{ij} associated with every arc, i.e., (i,j), are defined. These parameters represent the cost of shipping one unit from i to j directly along arc (i,j). The decision variables x_{ij} define how much of material to ship along each arc. In order to fine the minimum-cost, the objective is to minimize the total cost of moving the supply to meet the demand:

$$\min \sum_{(i,j) \in \mathcal{L}} c_{ij} x_{ij} \tag{3.8}$$

The flow conservation is satisfied by the following constraints. For each node $k \in \mathcal{V}$, the total flow into node k is given by:

$$\sum_{i:(i,k) \in \mathcal{L}} x_{ik}. \tag{3.9}$$

The total flow out from node k is:

$$\sum_{j:(k,j) \in \mathcal{L}} x_{kj} \tag{3.10}$$

The difference between these two quantities is the net inflow, which must be equal to the demand at the node b_k. Hence, the flow conservation constraints can be written as:

$$\sum_{i:(i,k)\in\mathcal{L}} x_{ik} - \sum_{j:(k,j)\in\mathcal{L}} x_{kj} = -b_k \qquad k \in \mathcal{V} \tag{3.11}$$

Finally, the flow on each arc must be non-negative[3]:

$$x_{ij} \geq 0 \qquad (i,j) \in \mathcal{L} \tag{3.12}$$

This formulation can be written in standard format of LP as:

$$\begin{aligned} \min \qquad & c^T x && (3.13)\\ \text{subject to :} \quad & Ax = -b && (3.14)\\ & x \geq 0 && (3.15) \end{aligned}$$

Which can be optimized using Simplex method. The solution of this formulation returns values of variables x_{ij} that define the flow distribution in network $G(\mathcal{V}, \mathcal{L})$.

If values of c_{ij} are equal for all links $\mathcal{L}$ then this formulation finds the *shortest-paths* in the network.

Maximal flow problem

The Maximal flow problem addresses when for a given network the maximal flow must be found. The foundation of this problem dates back to 50" and US Army which needed to find the best path in an military area in order to deliver as much supplies as possible. First, this problem was resolved by the *Ford and Fulkerson's labeling method* [36]. However, this method does not provide enough performance, so specialized primal simplex implementations was found. The simplex implementation is significantly faster than the original Ford and Fulkerson's labeling method[4], and requires only one-third of the amount of memory[35]. However, this method is a great algorithm for learning this category of problems. The maximal flow problem considers a network $G(\mathcal{V}, \mathcal{L})$ having $|\mathcal{V}|$ nodes and $|\mathcal{L}|$ arcs which support the commodity flow. Each arc (i, j) only supports flow with non-negative value of commodity, but restricted by the capacity of arc u_{ij}. Compared to the minimum-cost network problem this problem does not consider arcs to have a cost. Let f represent the amount of flow in the network from node s to node d. Then the maximal

[3]Negative value of variable x_{ij} represents opposite direction of flow.
[4]The speed up factor is 1.5-5.

flow problem may be mathematically formulated as follows:

$$\max \sum_{i,j \in \mathcal{V}} f \tag{3.16}$$

Subject to:

$$\sum_{i:(i,k)\in\mathcal{L}} x_{ik} - \sum_{j:(k,j)\in\mathcal{L}} x_{kj} = \begin{cases} f & \text{if } k = s \\ -f & \text{if } k = d \\ 0 & \text{otherwise} \end{cases} \qquad k \in \mathcal{V} \tag{3.17}$$

$$x_{ij} \leq u_{ij} \qquad (i,j) \in \mathcal{L} \tag{3.18}$$

$$x_{ij} \geq 0 \qquad (i,j) \in \mathcal{L} \tag{3.19}$$

$$\tag{3.20}$$

where the (3.17) represents the flow-constraints described in the previous section, and the constraints (3.18) create link maximal capacity. This is called the *node-arc formulation for the maximal flow problem.*

This formulation shows the maximal capacity that can be used by a flow $s \to d$ in given network. This formulation only allows one flow at the time.

Maximal throughput problem

Both of those previously mentioned methods provide a great background for optimizations using simplex method. However, both methods only allows one flow in the studied network. In the optical network which are studied in this thesis; however, it is very important to find routing for multiple flows at the same time because in OBS network sub-wavelength scheduling allows using one arc by multiple flows. The maximal throughput formulation uses the flow-constraint in order to find routing, but the link capacity constraint is more complex. The maximal throughput problem may be mathematically formulated as follows:

$$\min \sum_{s,d \in \mathcal{V}} n_{sd} \tag{3.21}$$

Subject to:

$$\sum_{i:(i,k)\in\mathcal{L}} x_{ik}^{sd} - \sum_{j:(k,j)\in\mathcal{L}} x_{kj}^{sd} = \begin{cases} b_{sd} & \text{if } k = s \\ -b_{sd} & \text{if } k = d \\ 0 & \text{otherwise} \end{cases} \qquad s,d \in \mathcal{V}, k \in \mathcal{V} \tag{3.22}$$

$$\sum_{s,d\in\mathcal{V}} x_{ij}^{sd} \leq u_{ij} \qquad (i,j) \in \mathcal{L} \tag{3.23}$$

$$0 \leq D_{sd} - b_{sd} \leq n_{sd} \qquad s,d \in \mathcal{V} \tag{3.24}$$

$$x_{ij}^{sd} \geq 0 \qquad (i,j) \in \mathcal{L}, s,d \in \mathcal{V} \tag{3.25}$$

$$b_{sd} \geq 0 \qquad s,d \in \mathcal{V} \tag{3.26}$$

where D_{sd} is an parameter defining the amount of required capacity between $s \to d$, and n_{sd} represents amount of unsupported traffic due to link capacity constraints (3.23), for example. The objective of this formulation is to support as much required capacity in a network as possible.

OO Optical networks also support multiple wavelengths which makes this formulation even more complex. Then it is very easy to fight with so called scalability when augmenting network about one link might exponentially increase the solution time. Thankfully, those wavelength does not share variables, so the constraint matrix A is sparse, so Dantzik decomposition is possible to use [32].

3.3 Methodology of mathematical model evaluation using simulations

It is very important to prove claims and improvements. Mathematical optimizations are a great tool, but as everything in real world, it can not be trusted without verification. Large-scale optimizations of routing in all-optical networks are great example.

When solving routing, the results give the optimal distribution of traffic across links of graph representing studied topology. In the first moment routing must be verified whether include loops or defects indicating problems in model formulation. At this point, it is always worth it to prepare a program that checks the solution validity. Results are simply transformed into forwarding instructions, and optical networks blindly follow them, so after a while in the production deployment a problem can be found unless all problems are found by verification of results.

Therefore, the implementation of simulations is very vital in order to improving mathematical model. As it has been mentioned in Section 3.1, simulations are cheap and very

good tool for broadening knowledge of systems. Moreover, simulations can provide information that would be obtained from a real deployment if the simulations model is well programmed. The performance parameters such as latencies, or probabilities, blocking, buffering, etc. which are caused by the shared media and are not possible to quantify using optimization models without detailed knowledge that simulations or real deployments give.

The scope of this thesis is to provide an insight into the performance of OBS networks. Using only simulations or optimizations exclusively does not lead to proper conclusions, but combining both methods is not trivial. In most of the cases, researchers usually stick with one of those methods and does not use the other one, or very little. Therefore, this section tackles a methodology of combining results of optimizations with verification using simulations. The workflow is depicted in Figure 3.6. The experiments in this thesis are carried out with Cplex (optimizations) [37] and OMNeT++ (simulations) [34]; however, both software does not share any libraries, and it is not possible to define output format which would allow transfer data from one to other. All this must be realized by hand. All those transfers are ensured by program written by Python language and universal data analyzer (UDA), see Section A.

The UDA is used for analysis of both the optimization and simulations. In case of simulations, every simulation is analysed separately, and saved into a binary format with proper indexing. The indexing is very important. When all simulations are finished, all these binary files are merged together in order to create a dataset which is then analysed using statistical methods. Indexing is then used in order to describe final dataset. Results of this statistical analysis with results from optimizations form the base which gives answer to the original problem. At the same time, the results can show that optimization model is incorrect, or simulation model is incorrect then it is necessary to repair models and start the experiment over.

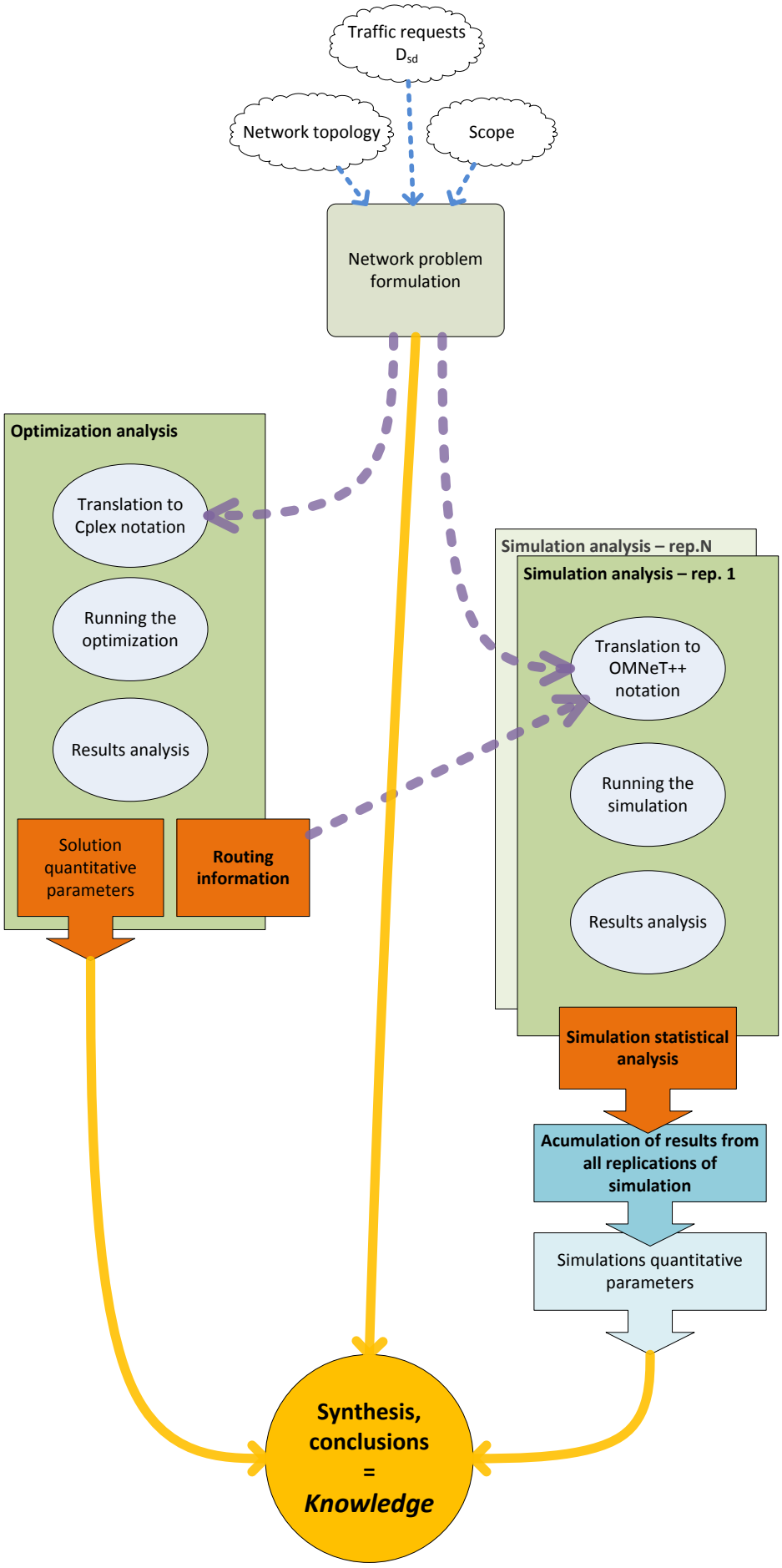


Figure 3.6: Visualisation of interaction between optimization and simulations based on an assumption that gets verified and synthesized.

Chapter 4

Optical Burst Switching

OBS [13], [38] is very promising architecture designed to transmit data traffic over an all-optical infrastructure. OBS was developed as an alternative to a low-flexible OCS and a technologically immature OPS solution. OBS has gained considerable attention in the past few years, and various solutions have been proposed and analysed. The scope has been to improve OBS performance, e.g., burst assembly techniques, channel scheduling schemes, contention resolution methods and QoS provisioning. Although still being developed, i.e., not standardized yet, OBS is considered the winning network technology to enable the next generation optical Internet [39]. Among other reasons, the success of OBS relies on two features: one-way signalling schemes, and small control overhead for a large amount of payload data. This chapter describes the main features of this technology, benefits, and challenges.

4.1 Classical OBS

An OBS network consists of a set of electronic edge and optical core nodes and nodes connected by WDM links, see Figure 1.4. Edge nodes aggregate data from client networks and assemble them into optical bursts. Each burst contains a data payload and is signalled by a burst header packet – BHP. The burst header packet is generated when the assembly process of the burst data payload is finished. The BHP carries all the information necessary in order to properly transfer the burst inside the network, so the burst is carried all-optically. Bursts and their BHPs do not share the same wavelength, but there is a physically dedicated channel for the network control. The control channels can be either out-of-fiber or in-fiber. The in-fiber method is used currently because network control and data transmission channels are impaired the same way. This is crucial for the scheduling because the inter-arrival time between BHP and burst must be the same as what is stored in BHP, in form of OT, otherwise signalling can not work properly.

The OT decreases everytime a BHP is processed by a core node, so the OT must be proportionally big to the number of hops otherwise burst would arrive before BHP, so the burst would not be possible to switch. In other words, the offset time is introduced in order to give time to both BHP processing and reconfiguration of the switching matrix. The core node controller performs several functions, among others the burst forwarding and resources reservation. Both functions, i.e., forwarding and reservation, work together in order to find the proper output port and proper time when this output port and its wavelength can be used. In case the wavelength is occupied by another burst a contention resolution mechanism, if implemented, is applied. The contention resolution mechanism may cooperate with a scheduling policy if alternative resources can be provided for the burst transmission. In case no resources are available for the incoming burst it is lost. After the burst transmission is finished in a node the connection in the switching matrix can be used for other connections.

4.1.1 Architecture

The OBS node function is realized by set of building blocks where each block does only a simple operation, see Figure 4.1. This figure addresses function of nodes by their logical location that means Edge node must provide functionality as it was an egress and ingress node.

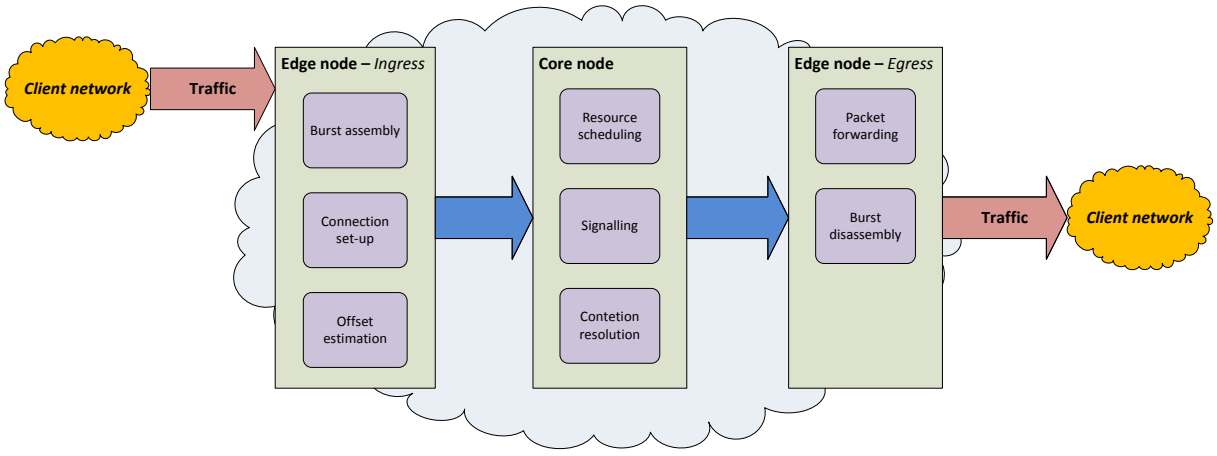


Figure 4.1: Logical distribution of functionality that nodes must provide in order to support the OBS network functionality.

Core node

Core nodes are located in the optical part of OBS network where all-optical switching takes place, see Figure 4.1. OBS core nodes perform two types of operations a) all-optical

switching b) BHP processing. Based on the information in BHP and scheduler, the control processor can either schedule switching or can do contention resolution. This function is enabled by the Core node architecture that is captured in Figure 4.2.

OBS networks use WDM in order to use the fibers as efficiently as possible. In order to decouple the control channel, which sits on a dedicated wavelength, WDM demultiplexers are used as the first step. Demultiplexers forward the dedicated wavelength towards the *Electronic switching controller* and other wavelengths towards the *MX*, see Section 1.6.1. The electronic switching controller is the brain of Core node because it realizes the computational functionality: forwarding, scheduling, and contention resolution. In order to do that, the electronic switching controller contains memory where information about arrival times and burst durations are stored. Records in this memory are used for scheduling, and in case there is no available wavelength at the time burst arrives the contention resolution mechanism is initiated.

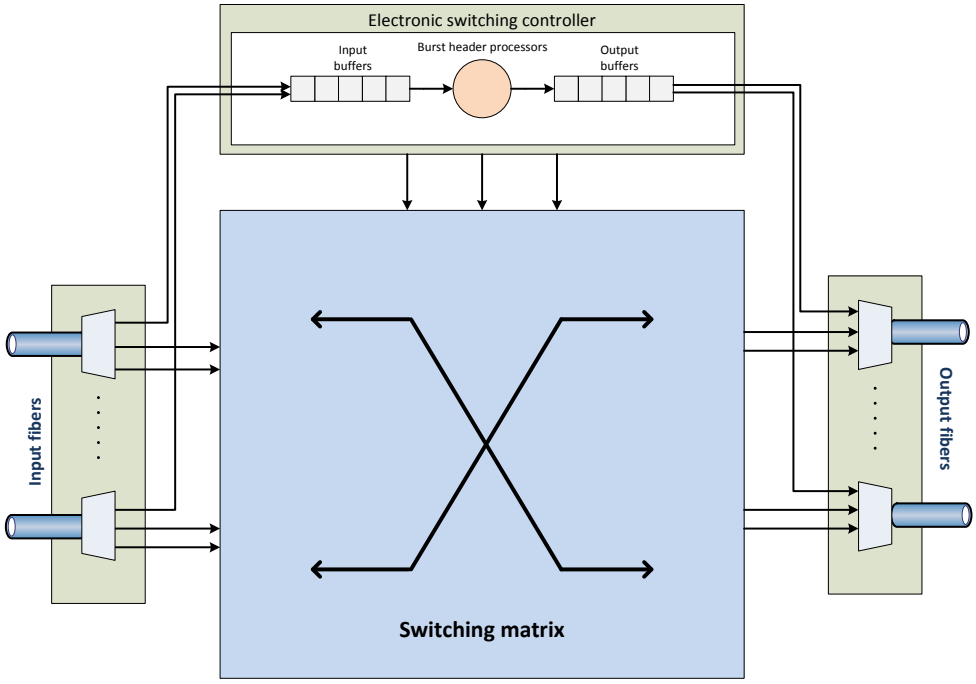


Figure 4.2: Core node architecture (Figure adapted from [40]).

OBS Core node is very simple network node which simply follows instructions that come to the node according to the signalling, see Section 4.1.2.

Edge node

Every edge node acts in two very different positions a) Ingress node b) Egress node, see Figure 4.1. These two positions require mutually exclusive set of function blocks, but from the logical perspective, these two sets of function blocks are installed into one network

node, i.e., receiver and transceiver side. Ingress node provides: *burst assembly, connection setup*, and *offset estimation*, and the egress node provides *burst disassembly* and *packet forwarding*. The general Edge node architecture is captured in Figure 4.3

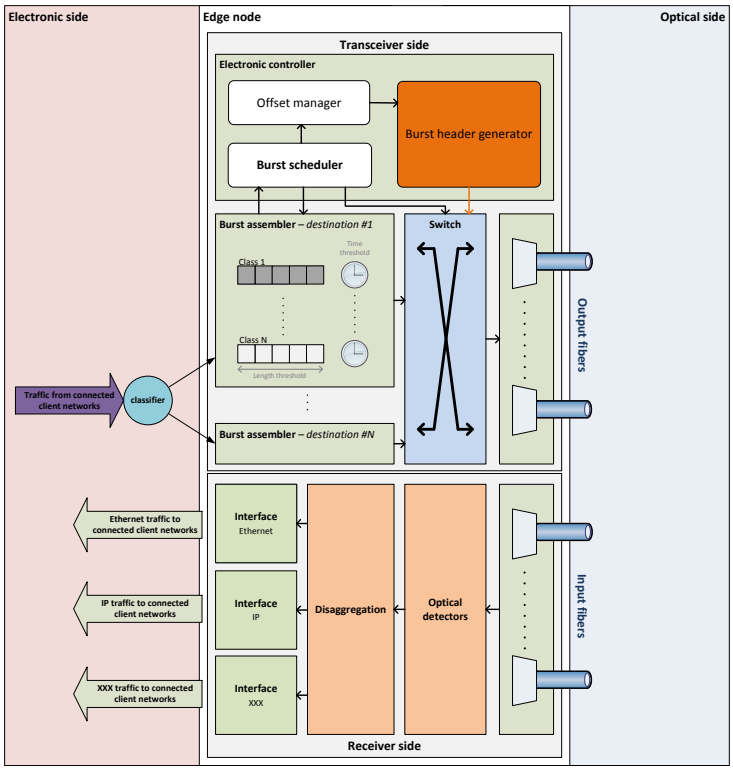


Figure 4.3: Edge node architecture (Figure adapted from [40]).

Burst assembly is the process of assembling data packets are assembled into bursts. Each edge node maintains multiple queues to aggregate data packets from client networks according to their destinations and QoS requirements. When some of these queues are full according the an assembly algorithm, i.e., time and length trigger, QoS, or more sophisticated triggers based on traffic properties, a burst is created from these data packets. The timer is used by the edge node to decide when to assemble and initiate a new burst. The minimum and maximum burst size parameters determine the length of the assembled burst. These parameters require careful setting because long bursts can hold network resources for long time periods; subsequently the burst loss can increase. On the other hand, short bursts result into the increase of number of BHP, more often switching, and requirement on faster MX, but allows smoother grooming.

Connection setup is an process ensuring all criteria for successful transmission are met. Connection can only be established if switching resources are allocated and configured along the path of the transmission. This composes of three components i) signalling ii) routing iii) wavelength assignment.

Offset estimation is the perennial process which makes the difference between OBS and OPS. The transmission of every burst is preceded by the transmission of a BHP, see Figure 4.4. More precisely, after sending the control packet, the edge node waits for a fixed or variable offset time until it starts transmitting the corresponding burst. The OT is used to allow the BHP processing, core node reserved the required resources, and core node configured the MX. If the BHP succeeds in resource reservation then the arriving burst can be bypassed at each intermediate node without any buffering or processing. Therefore, an appropriate estimation of the OT is crucial for the network performance.

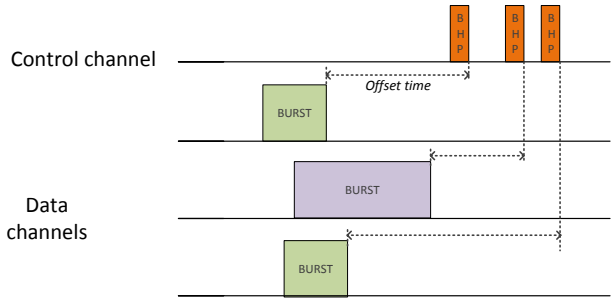


Figure 4.4: Edge node architecture (Figure adapted from [40]).

Burst disassembly is very simple process which is preceded by the detection of optical signal, i.e., O/E conversion. The input of the burst disassembly is burst in electronic format. This electronic burst is processed by Field Programmable Gate Array (FPGA) in order to withdraw frames, packets, or etc. in the proper format because on burst can transfer different protocols at the same time.

Packet forwarding represents the actual edge of OBS network. An Edge node can contain a number of packet forwarders because each is dedicated for different client network technology. Full IP stack with routing table is necessary for IP networks, for example. Based on the information in the packet forwarder the payload is forwarded to the corresponding port which represents a client network.

4.1.2 Signalling

Signalling is the perennial component of every network because it creates paths that ensure the transmissions itself. This is very important in OBS networks because without proper signalling bursts will never reach the right destination. Signalling specifies the protocol that handles the transmission and how efficiently a network can be used. Several variations of signalling protocols exist for OBS depending on how and when the resources along a route are reserved for a burst. Several variations of OBS signalling protocols are possible, depending on how and when the resources along a route are reserved for a burst.

In particular, a signalling scheme can be characterized by the following characteristics as captured in Figure 4.5.

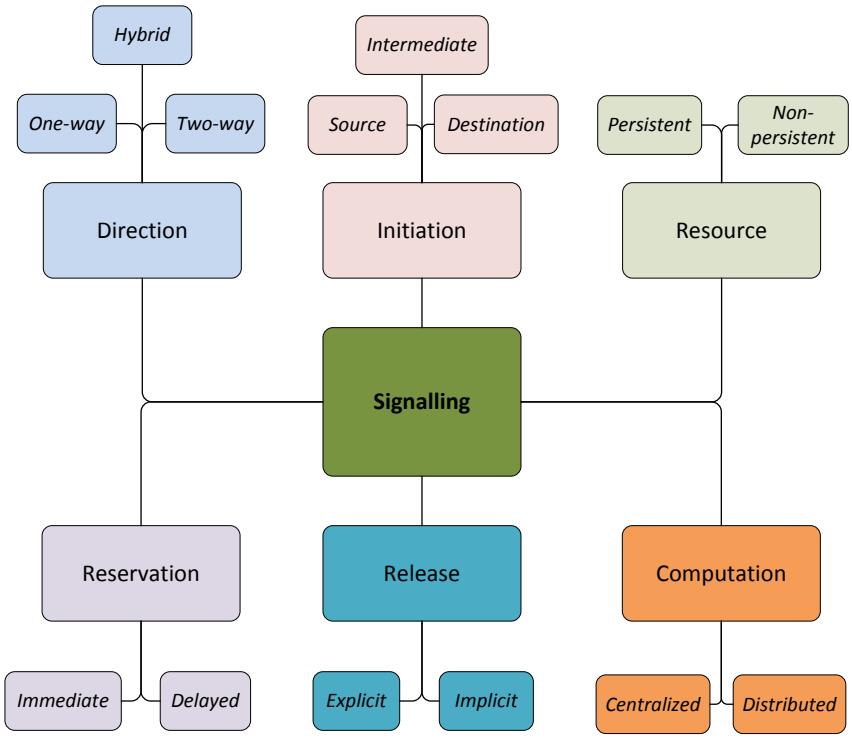


Figure 4.5: Signalling mechanisms in OBS (Figure adapted from [41]).

Direction – Signalling scheme with *one-way* reservations is initiated by the source node when a node sends out a BHP requesting each node along the path to allocate the necessary resources for the data burst and to configure their MX accordingly. The source node then sends out the burst without waiting for an acknowledgement from either the intermediate nodes or the destination node regarding the success or failure of the resource reservation process at each node. On the other *two-way* signalling scheme always ways for acknowledgement from either the intermediate nodes or the destination node then sends out burst. In *hybrid* signalling schemes, the reservations from the source to intermediate node in the route are confirmed through acknowledgements, while the reservations from the destination are unconfirmed. One-way signalling allows reaching lower end-to-end delay compared to two-way signalling scheme.

Initiation – A signalling technique can initiate the reservation of the requested resources at the source, destination, or intermediate hop. In case of *source initiation* a burst is send when acknowledgement from all intermediate nodes confirm selected wavelength which was proposed by BHP. On the other hand in *destination initiation*

burst is only send when acknowledgement from destination node arrives. This acknowledgement also carries information about selected wavelength. For the *intermediate initiation* some intermediate nodes confirm to the source node and BHP collects information about available wavelength, and destination node sends information about possible wavelengths to the source node.

Resource – represents the elementary MX connection which can be used at the moment when it is requested for reservation. In case the resource is reserved *persistently*, if it is requested, then the resource is indicated as blocked for this traffic. On the other hand, if the resource is *not* used *persistently* then the resource can be reserved, but if there is another request for the same resource contention can occur and right connection failure mechanisms is initiated, i.e., re-transmission, deflection, and buffering.

Reservation – can be realized immediately or later when a BHP is processed by Electronic switching controller of core node. Signalling with delayed reservation allows achieving higher wavelength efficiency, but it is more difficult to implement, e.g., just-enough-time (JET). Signalling with immediate reservation is simple and practical to implement, but incurs higher blocking due to inefficient bandwidth allocation, e.g., just-in-time (JIT).

Release – In an *explicit release* technique, a special control message is sent along the path after the burst is sent in order to indicate that burst transmission was finished. The core nodes then can use the elementary connection in MX use again. On the other hand, in an *implicit release* technique, the BHP carries additional information such as the burst length and the OT because then the elementary connection in MX can be release time can be calculated by the core node.

Computation – defines the way the signalling is distributed in OBS network. When there is a central server which reserves wavelengths in network and in time then this approach is call *centralized signalling*. In *distributed signalling* techniques, each node has a burst scheduler assigning an outgoing channel for each arriving BHP in a distributed manner.

The main focus of this thesis will by on one-way, source imitated, non-persistent, delayed, distributed, and implicitly released signalling which is generally recognized under name JET. Similarly to the JET, there is a signalling scheme called JIT which differ only in the way the reserve the resource.

Just enough time – JET

JET signaling makes use of the OT information transmitted in each BHP. JET allows higher wavelength utilization due to its delayed reservation feature. With delayed reservation the optical switches at a given OBS node are configured right before the expected arrival time of the data burst (and not immediately after the control packet, like in JIT). When a burst is ready for transmission the JET signalling scheme works as follows the source node sends a BHP toward the destination node; the control packet is processed at each subsequent node in order to establish an all-optical path for the corresponding data burst; if this reservation attempt succeeds, the switches along the path will be properly configured prior to the burst arrival, see Figure 4.6. In the mean time the burst waits at the source, in the electrical memory of burst assembler. Upon expiring a predetermined offset time, the burst is send optically on the selected wavelength.

The OT is calculated based on the number of hops from source to destination and the switch reconfiguration time of a core node in the following way:

$$OT = h.d_p + d_s, \tag{4.1}$$

where h is the number of hops between source and destination, d_p is the per-node BHP processing time, and d_s is the switching reconfiguration time [41]. If at any intermediate node the reservation is unsuccessful the burst is dropped.

4.1.3 Scheduling

A scheduling algorithm is responsible for a decision which wavelength is going to be assigned by a given burst in case there are more wavelengths available. The simplest scheduling schemes can be based on either a random or a round-robin wavelength selection. However, bursts may have different sizes, have different OTs, and not arrive in the same order as their corresponding BHP hence channel occupancy becomes fragmented with different voids between bursts. In order to achieve high efficiency of wavelengths, an scheduling algorithm must be able to use even those voids between bursts. Scheduling algorithms can be roughly classified into two categories[41] i) non-void-filling scheduling algorithms ii) void-filling scheduling algorithms. The non-void-filling scheduling algorithms are easier to implement, but they are less efficient in terms of wavelength usage. On the other hand, the void-filling scheduling algorithms are more difficult to implement and require more memory in order to store information about scheduled burst, but the void-filling scheduling algorithms provide higher level of wavelength usage.

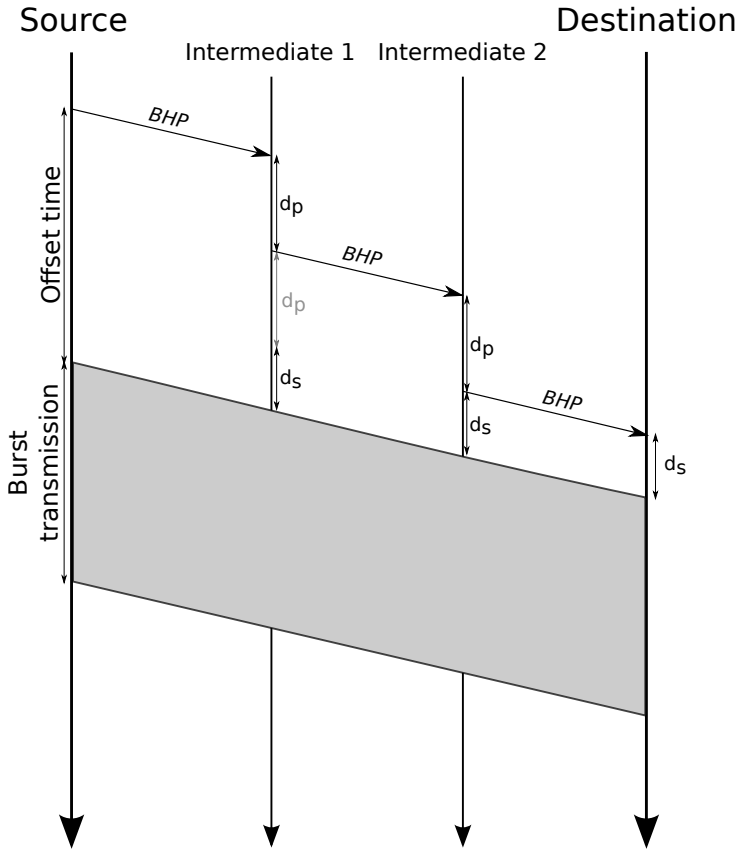


Figure 4.6: JET signalling scheme (Figure adapted from [41]).

Latest available unused channel (LAUC)

LAUC is a Horizon-type algorithm. It means the scheduler keeps track of the so-called horizon of each wavelength channel. Channel horizon is defined as the latest time at which the wavelength is scheduled to be in use, see Figure 4.7: Channel 1 – t_2, Channel 2 – t_3, and Channel 3 – t_5. Only the channels whose scheduling horizons precede the arrival time of the new burst are considered available, and the one with the least waiting is chosen. After choosing this channel, the horizon of this channel is update to the time when the new burst sending finishes. The LAUC scheduling algorithm would chose the channel 2 in situation as depicted in Figure 4.7 because Channel 3 horizon is t_5.

The weakness of this algorithm is that it can not use voids which leads to low bandwidth utilization and high burst loss probability.

LAUC-VF

LAUC-VF is a JET-based algorithm. It means the scheduler keeps track of all existent voids in all wavelength channels. This algorithm assigns an arriving burst which is large enough to fill a void interval whose starting time is the latest but still earlier than the burst arrival time. The LAUC-VF scheduling algorithm would chose the channel 3 in

situation as depicted in Figure 4.7.

○○ This algorithm provides better bandwidth utilization and burst loss probability than LAUC. The complexity of this algorithm is much higher than LAUC.

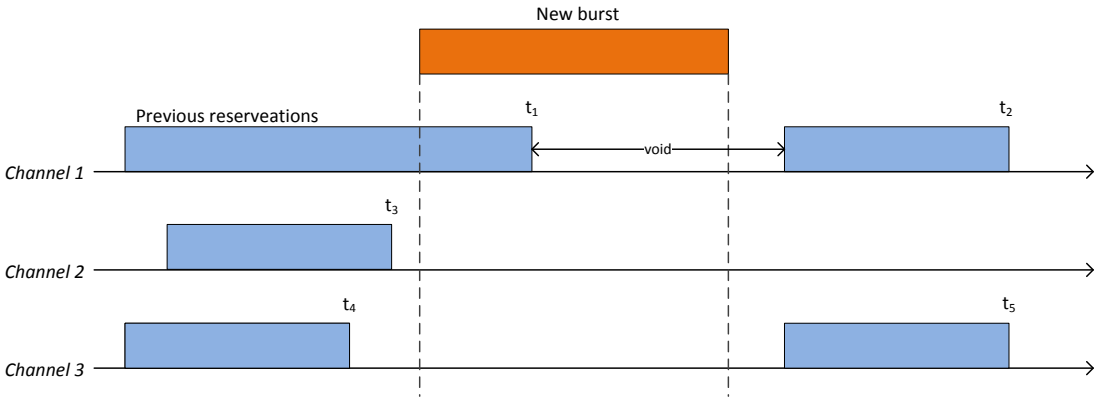


Figure 4.7: Wavelength selection under two similar scheduling approaches. (Figure adapted from [41]).

4.1.4 Contentions

Contentions are the perennial weakness of OBS networks because can cause burst loss, and the burst loss caused small acceptance by the research community. In this chapter I am going to focus on sources of contention, and techniques mitigating effects of contention. In the section 4.1.2 types of signallization in OBS networks were recalled. Currently, OBS networks use JET signalling, see Section 4.1.2, that is very bandwidth efficient, one-way signalling reservation scheme without channel acknowledgement. The problem of the JET is the shared medium that is used for transmission of some flows at one wavelength. Therefore, if a burst arrives at an OBS node and all local resources are already taken, or if two or more simultaneously arriving bursts compete for the same resource then the contention occurs. This situation might lead into burst loss. Burst loss is a critical issue with significant impact on any end-to-end application running in the upper layers because it reduces overall performance. Contention is the only[1] source of burst loss, so contention resolution schemes are the main design objectives in OBS networks and they have been deeply studied [26]. The general resolution of burst contention is to move all but one burst out of the way. In OBS networks, this can be accomplished by actions in the time, space or wavelength domains, or any combination of these actions. Examples of contention resolutions actions are: optical buffering (time domain), deflection routing (space domain), and wavelength conversion (wavelength domain).

[1]In terms of transmission system. Fiber cut causing rerouting and some burst loss is not considered in this statement.

Optical buffering

Fully optical RAM memory still is not ready for production use. The only meaningful way how to buffer light is FDL which only gives fixed delay, but cannot store the light for variable time duration. Optical buffers are either constructed in single stage, which have only one block of delay lines, or in multiple stages which have several blocks of delay lines cascaded together, where each block consists of a set of parallel delay lines. Usually, FDLs are connected to some ports of MX, and any input port can be switched into the FDL. The output of FDL is connected on a input port of MX, so the light can be buffered a number of times.

The main disadvantage of FDL are the dimensions. If one considers that 200 km of optical fiber gives delay 1 ms, and more FDLs giving various delays are necessary then this system can not be used in the outlayed areas.

Wavelength conversion

Wavelength conversion can resolve contention by putting the contenting burst onto different wavelengths where they do not content, and these bursts are send by the same output link. Wavelength conversion helps to increase the wavelength reuse and the wavelengths can be spatially reused to carry different bursts on different fiber links in the network. Wavelength conversion is of four types: full wavelength conversion, limited conversion, fixed conversion, and partial wavelength conversion. *Full conversion* allows any incoming wavelength to convert to any outgoing wavelength; whereas, the *limited conversion* allows conversion of some incoming wavelengths to be converted to all the outgoing wavelengths [41]. *Fixed conversion* allows conversion of each incoming wavelength to one or more predetermined wavelengths only. In *partial wavelength* conversion some nodes can do some types of wavelength conversion in network.

Deflection routing

Deflection routing technique allows deflection of some contenting bursts to be switched onto alternate paths toward the destination in case. Deflection routing requires changing the offset time of a burst which is impossible without the use of buffers. Subsequently, only limited number alternative paths is possible.

Burst segmentation

Burst segmentation can be seen as a trade-off between burst loss and transmission. burst segmentation allows to drop only the overlapping part of the burst. Therefore segmentation can be classified into head dropping or tail dropping depending which part of

burst is dropped. The remaining segment of the burst is transmitted successfully to the destination thereby increasing the packet delivery ratio.

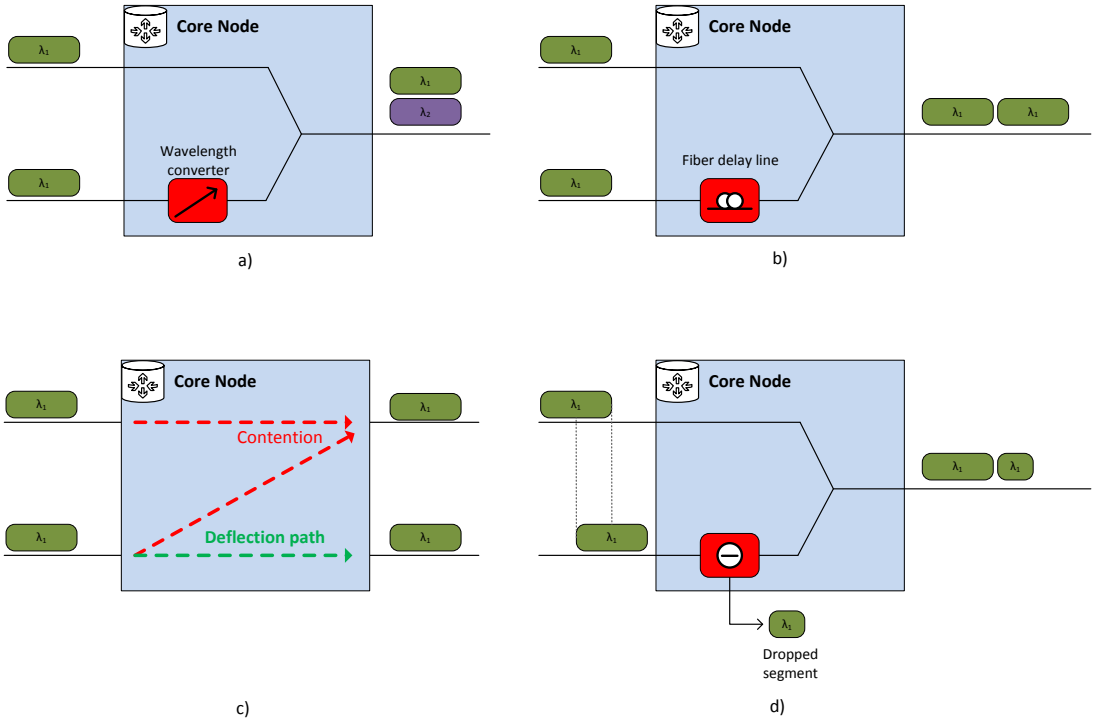


Figure 4.8: Visualisation of basic contention resolution schemes. (Figure adapted from [17]).

4.1.5 Stream-line effect

One of the unique features of OBS networks is that there is no buffering inside the network. Although FDLs may be used, they can only provide limited and deterministic delay. It implies that any burst loss is caused by wavelength contention instead of buffer overflow. Traditionally, the M/M/k/k queueing model is adopted in performance evaluation of OBS networks. This model assumes that the input traffic to an OBS core node is Poisson, which is equivalent to having an infinite number of independent input streams. However, the number of input streams to a core node is bounded by the small number of input links. Therefore, the actual burst loss probability at an OBS link is smaller than that obtained from the M/M/k/k model and is strongly dependent on the number of input streams and their relative rates. It is caused by the fact that bursts within one input stream are streamlined and only inter-stream contentions happen at the link. This phenomenon is called the *streamline effect*, and is depicted in Figure 4.9.

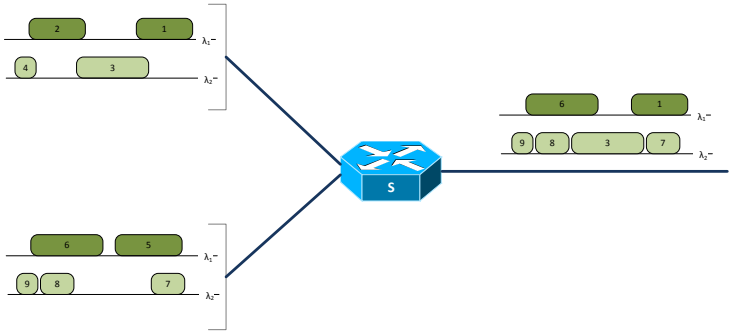


Figure 4.9: Illustration of stream line effect. It is clear that only bursts 2, 6 and 4, 9 are contenting and other do not at the left side of the node. On the output side no burst are contenting.

4.2 Loss-less OBS architectures

The previous section tackled the biggest issue of OBS networks, contentions. The contentions are the main reason why OBS networks are not preferred. The disadvantage of the aforementioned contention resolution approaches is that they increase control or architecture complexity due to necessity of additional systems, e.g., wavelength converters, FDL, to be installed. Therefore, after some time of stagnation in the research in the field of OBS networks a new trend started to be seen, i.e., Loss-less OBS.

The objective of loss-less OBS networks is to allow no loss. In other words, loss-less is only possible if there is a way how to mitigate contentions. Contentions can be mitigated, but under some price. The price of loss-less depends on the proposed architecture. The price can be seen as increased end-to-end delay, lower wavelength efficiency, etc. In the following text some architectures are described in order to give overview of possibilities.

4.2.1 Packet Optical Add/Drop Multiplexing

Packet Optical Add/Drop Multiplexing (POADM) [42], [43] is a ring burst-switched solution using synchronous time-slots for bandwidth allocation. It is developed in collaboration between NTT Photonics Labs and Alcatel-Lucent Bell Labs. It is originally designed to be deployed in a WDM ring network operating at 10G with different modulation formats over 40-wavelengths.

Data are transported in slots with fixed duration 10 μs and can transport only one burst. Moreover, each time-slot includes one guard band, one preamble, one synchronization word and a payload. Dummy bursts are generated when no traffic is sent in the network to simplify the power management in the amplification stages. A dedicated channel is used for network control. This control channel is put onto a separate wavelength. POADM networks are designed using two kinds of elements a) Hybrid Optoelectronic

b) Packet Router (HOPR) [42]. HOPR ensures the interconnection between POADM rings in the metro network architecture. When bursts transit between rings, bursts are switched by single or several HOPRs. Every time this switching does not cause contention bypass is used; otherwise the bursts are electronically buffered and switched to the output port when they cannot cause any contention. Due to ring topology, i.e., no flow merging, contention can not occur as indicated by SLE [44].

⚯ POADM uses electronic buffering and ring topology in order to mitigate contentions.

4.2.2 Optical Burst Transport Network

Optical Burst Transport Network (OBTN) [45] is an all-optical transport network architecture proposed by Huawei. OBTN is a time-slotted architecture based on an out-of-band and centralized control plane. The control wavelength carries configuration and slot reservation information from central control entity to the local control unit of each node.

Each wavelength is attributed to each source to transmit data to the other nodes of the network. Each wavelength is divided into equal time slots, called Optical Burst (OB) slots. OBs are grouped into frames and are time aligned with the other OBs on the other wavelengths. Two OBs occupying the same slot of time on two different wavelengths should not be destined to the same node. Accordingly, nodes are connected to each other by disjoint OB virtual paths.

⚯ Contention can not occur due to non time overlapping virtual paths from different sources to one destination.

4.2.3 Wavelength-Routed Optical Burst Switching

Wavelength-Routed Optical Burst Switching (WR-OBS) [46] combines OBS with dynamic wavelength allocation under fast circuit switching. WR-OBS might be considered to be closer to dynamic circuit switching since the transmission of a burst between two edge nodes requires a dynamic set up of an end-to-end lightpath. The lightpath establishment process is based on two-way reservation mechanism between the edge node and a centralized control entity. More precisely, client layer's packets are aggregated in the edge routers into bursts according to destination. When burst meets assembly triggers, the edge node sends wavelength request to the control node to transmit the burst. Once the RWA finds an available free wavelength, the control node sends acknowledgement to the source to emit the bursts and it sends also several control messages to intermediate nodes to configure their MX.

⚯ Contention can not occur due to centralized network management. Flow-paths are

actually only very fast lightpaths, so all principles related to OCS apply onto WR-OBS as well.

4.2.4 Optical Packet Switching and Transport

Optical Packet Switching and Transport (OPST) [29], [47] is proposed by Intune Networks, aims to create a robust, and asynchronous packet-switching architecture in the form of a distributed non-blocking Ethernet switch. This approach intends to be deployed in the MAN.

Data are transported by two contra-rotating optical rings forming two autonomous and redundant packet switching fabrics. The burst transmission system is designed to balance traffic across these two optical switch fabric planes. OPST is based on wavelength routing scheme to address packet flows. The transmitter is equipped with a Fast Tunable Laser (FTL) to rapidly switch wavelength according to the target destination; whereas, the receiver has a fixed wavelength filter, thereby the wavelength acts as the address. The OPST network is composed of a set of parallel shared media. The access to a specific media relies on an Optical Media Access Control scheme inspired by the way Carrier Sense Multiple Access with Collision Avoidance (CSMA-CA) avoids collisions. When a burst is composed, the laser looks for a gap in the optical spectrum to transmit data. To do this, each node is equipped with an optical sensor allowing finding the channel occupancy. When the transmission system detects that the channel is free, the burst is inserted. When an upstream optical signal is detected, the burst emission is interrupted and it is resumed as soon as the channel returns free. This function is enabled by additional FDL which allows transmitter to stop sending data without any collision.

○○ Contentions are not possible due to ring topology, but collision are still possible due to asynchronous transmission mode. In order to avoid collision, CSMA-CA modification is used in order to detect transmission on shared medium.

4.2.5 CAROBS

CAROBS is a research project defined by Coutelen in his dissertation thesis [48], [49]. The aim of CAROBS is to create a robust, asynchronous, and all-optical mesh network. The robustness and efficiency is achieved with a new transmission mechanism called *burst train* which allows grooming in the optical domain and loss-less mode of operation due to contention resolution using electrical buffering.

CAROBS is a relatively controversy framework because it is labelled as all-optical framework, but uses electrical buffering. Buffering is a very promising approach of contention resolution in OBS [50], but for the time being, electrical buffering is the only

method of buffering for a variable time in optical network currently because purely optical memories are not ready for commercial deployment and FDLs are space inefficient.

The node architecture defined by CAROBS framework unifies OBS Core and Edge into one CAROBS node architecture which serves both purposes. The unified architecture is beneficial because only this way an edge node can sense traffic, i.e., flows, between ports. This is very important for collision avoidance, see OPST. The CAROBS node architecture is captured in Figure 4.10. The Figure 4.10 contains visualisation of burst train, cars, traffic aggregation and disaggregation, and logical process of buffering into media access control (MAC) block that maintains scheduled burst trains in its internal memory. The CAROBS node architecture depicted in Figure 4.10 contains horizontal lines representing the input and output ports. Additionally, the ports are labelled with λ to emphasize different wavelengths. The λ_1 represents the wavelength reserved for the control channel and λ_x represents the wavelengths used for optical transmission. In Fig. 4.10, one can see that the CAROBS node architecture spans three logical layers, with the control plane on the top containing an *SOA Manager* that reads CHP and determines further node actions. According to the CHP's content, the CAROBS node can: a) switch a whole burst train; b) groom-out a first burst and switch the rest; c) buffer contending cars to the electrical memory. In order to do this, the *SOA Manager* creates a set of instructions for the MX, then forwards the CHP to the next node on the dedicated wavelength λ_1. The central part is the electrical domain, and it is used for both the aggregation and disaggregation of user traffic as well as buffering contending burst trains in the MAC block. The bottom layer strictly resides in the optical domain. This layer is analogue only and represents the optical domain, where all cars are switched, and contains the MX block that switches the optical signal. If contention occurs, the contending burst trains are switched to the dedicated port for electrical buffering. These ports are directly connected to O/E blocks which convert optical signals to the electrical domain.

This function is enabled by the transmission mechanism called burst trains. A burst train is depicted in Figure 4.11. Every burst train is composed of a number of bursts (in CAROBS terminology called cars) which are located along the same flow-path, signalled by one header, and individual cars are separated by the time necessary for switching d_s between them. The void time between two consecutive cars allows a head drop which realizes all-optical grooming. The d_s time is equal to the time necessary for MX port reconfiguration in order to transmit an burst train from input port into its output port. Cars in an burst train are sorted according to the their OT length that represents the number of hops along the flow-path from source to destination. Every burst train tends to use requested time as efficiently as possible, so the intermediate time between two consecutive burst is minimized, but this minimization prone to OT enlargement hence

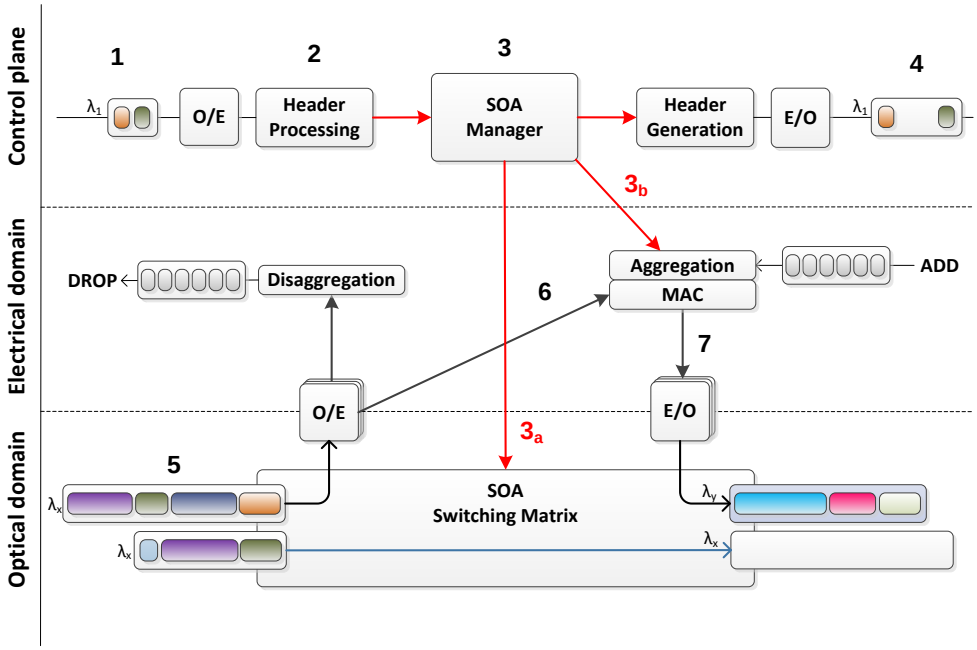


Figure 4.10: The CAROBS node resolves burst contention using an electrical buffer. The contending burst train is sent to the MAC and stored until the output port and wavelength are available again.

the OT extension is denoted by extra offset time (EOT) that is a positive number. Every burst train is signalled by one CHP which contains two set of information divided into a) train b) car section. The *train section* is used for routing and *car section* is used for grooming. A car is groomed out when burst train reaches its head car's destination, so head drop is carried out. The CHP structure is depicted in [48] Figure 8.3.

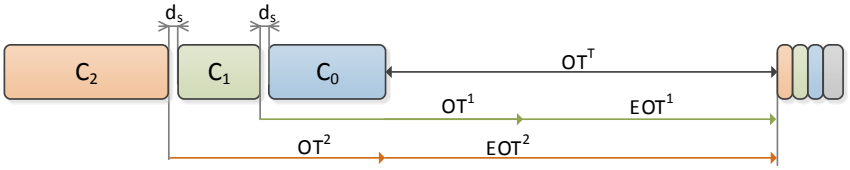


Figure 4.11: An example of burst train structure which contains a CAROBS Header and sequence of cars transferring clients data.

CAROBS framework does not define a way traffic is routed, i.e., RWA. This is the main contribution of this thesis. Nevertheless, the CAROBS framework was verified with shortest-path approach, and showed considerable improvement compared to OBS [49]. When RWA is carried out, every source node gets one path for each destination. Burst train contains only paths that overlap as depicted in Figure 4.12. Each path is used by one burst that is created from an AQ which can be seen as a queue where packets/frames from client network are stored prior they are send through the all-optical network. Data stored in AQ are changed into a burst only when time or space trigger is initiated. Space trigger represents the maximal time that an AQ can wait to be released, and the space

trigger represents the maximal size of burst. Both trigger must be carefully set because they very significantly influence the network performance.

Every AQ is part of a number of aggregation pools (APs). AP represents a virtual combination of AQ which bursts are routed along the same flow-path, see Figure 4.12, and each AQ can be head of only one AP. This means when a AQ is triggered by either trigger only one AP is initiated, and this initiation of an AP leads to initiation of a number of AQs that are destined along the flow-path.

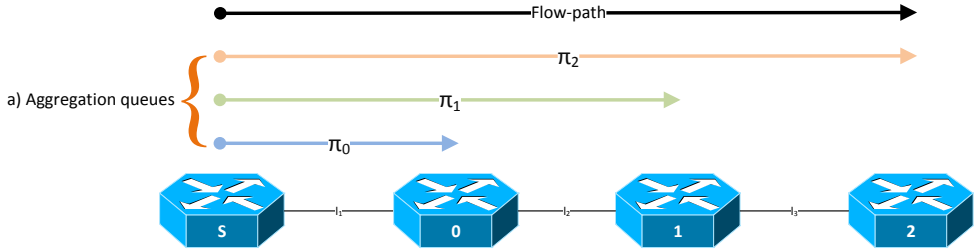


Figure 4.12: Visualisation of individual flow-paths for each destination. Each flow is labelled by π_x, see that colours of each path is equivalent to colours used in the description of burst train, see Figure 4.11.

Buffering

The main devise of CAROBS framework is contention resolution tackled by buffering in electrical domain. Several concepts using electrical memory for contention resolution have already been proposed [51], [52], but these concepts are too complex. The unique CAROBS node architecture allows sharing most of electronic resources necessary for electronic edge and the buffering in core part of a network. The buffering in the core part of the network starts when a new burst train is about to arrive, the CHP arrives to CAROBS node **1** in advance and on a dedicated wavelength. Subsequently, the CAROBS header undergoes optical detection (O/E) where the optical signal is adjusted and equalized. Following this, the control data are prepared for control processor **2**. Control processor **3** verifies the time span needed for the reservation of the burst train. In case of contention, the requested time span overlaps with already-scheduled burst trains, therefore, the SOA manager instructs MX and MAC **3a,b** to buffer the contending burst train. The SOA manager uses one instruction for MAC to set the buffering delay d_b, and two instructions for MX – buffering and unbuffering. These two instructions are assigned to MX with respect to the OT and JET signal protocol, so MX activates the configurations in $OT - d_s$ and $OT + d_b - d_s$ respectively.

This efficient process allows the CAROBS framework to ensure zero burst loss. However, the CAROBS framework breaks the original OBS paradigm [13] and increases the delay in the optical segment of the network. Yet the CAROBS framework increases per-

formance considerably [49] and provides the loss-less mode which is the central reason to adapt it further.

Chapter 5

Properties of CAROBS networks

The RWA tackles the problem of optical network dimensioning in terms of allocating bandwidth onto specific wavelength while holding optical network constraints. This means capacity of wavelength can not be over assigned by requested capacity of flows for every link in network, maximum of requested traffic must be transported, all the traffic must be link disjoint at one wavelength (OCS), traffic can be merged electronically (OCS), or etc. Very deep review of routing in OBS network was carried out by Klinkowsky *et al.* [26]. Klinkowsky *et al.* reviewed around 40 research papers dealing with RWA, categorized them according to their properties (Class, Routing type, Route selection, Features), route calculation approach, and suggested an optimization formulation based on LP considering M/M/k/k model.

This is very traditional approach, but it is very unfortunate because this approach does not consider SLE, see Section 4.1.5, [44]. Additionally, the CAROBS brings few features that must be considered by RWA as well hence this section opens the problem of CAROBS RWA formulation with description of statistical properties of CAROBS traffic, and the following sections going to tackle the CAROBS RWA dimensioning perspective and related problems. The main focus of the following analysis are to use already given optical network as efficiently as possible.

5.1 Statistical properties of CAROBS traffic

In the very beginning of research of CAROBS network the main focus was on the technical viability of this concept. The quantification of the number of O/E blocks which are necessary in order to deliver loss-less mode of operation was the main objective. It, however, turned out that this type of CAPEX and OPEX analysis are not viable unless the traffic properties are quantified. This observation obtained from experiments with online RWA analysis using shortest-paths and dynamic wavelength assignment [53].

The CAROBS architecture gives new traffic characteristics that are related to OBS traffic characteristics. An insight into the OBS traffic characteristics were carried out by Pavon-Marin *et al.* [50], and they concluded that buffer-less OBS architectures are not viable for mesh topologies. According to [50], the limiting factors of OBS are an inter-burst gap, a separate control wavelength and optical contention resolution. Therefore, the only way how to overcome the OCS performance is to change the OBS paradigm [50]. The separate control channel and inter-burst gap cannot be omitted unless a time-slotted approach is used [54]. Therefore, buffering seems to be the only way to increase OBS network performance in optical mesh networks. Unfortunately, very little work has been done on OBS buffering. Some early concepts used FDL [55], and subsequent papers deal with dimensioning FDLs [56]. Utilizing lengthy fibers, however, is problematic on most premises; therefore, some authors investigated scenarios utilizing electrical memories [49], [51]. Among these works [49], [51], [52], there is the CAROBS framework proposed by Coutelen *et al.* [49], see Section 4.2.5, but none of these works quantified the traffic characteristics. Traditionally, the traffic characteristics of OBS networks has been quantified using M/M/k/k models [40], [57]. However, this analytical model is not accurate, as it does not reflect the SLE [44], see Section 4.1.5.

The bottom line is, the majority of the works, which have been written, do not properly quantify OBS traffic characteristics because these works neglect the SLE. Additionally, the CAROBS framework uses buffering for contention resolution which means the buffered traffic must be restored when it is possible. However, this un-buffering can cause contentions which are called *secondary contention*. The secondary contention occurs when a burst train is scheduled to be un-buffered from electrical memory back to the optical domain on a given wavelength at the same time as an incoming burst is requesting the same wavelength. In such a case, the arriving burst train is buffered. This means neither the M/M/k/k , nor the SLE model work for buffered OBS exclusively. Secondary contention was studied by Delesques *et al.* [58] using the Engset model. Their main concern was the buffer size dimensioning, so they defined BP. BP is a comprehensive parameter of buffered OBS networks hence cannot be easily quantified by M/M/k/k model, with respect to SLE. Therefore, the SLE model and secondary contention must be mutually evaluated.

5.1.1 CAROBS traffic model

The CAROBS traffic characteristics are closely bounded to the buffering process that has implications into the CAROBS node dimensioning problem. At the first glance, the optimization of the node architecture might seem to be easy using the Erlang C formula. The Erlang C formula gives burst blocking probability (BBP) that seems to be equal to

the BP. However, there are two perennial shortcomings: the Erlang C formula only works for systems with buffering before the service [33]. Moreover, the contenting burst cannot be buffered unless there are enough O/E blocks. It means that there are servers, which need to be optimized, before the buffering can be optimized. Therefore, for the purpose of further discussion the CAROBS node architecture must be reformulated using the tools of Queueing theory [33]. The CAROBS node architecture then changes from Figure 4.10 to Figure 5.1. Such a redefinition is vital because it allows separation of two optimization problems a) the buffering problem itself b) optimization of the number of O/E blocks. Due to this separation these two problems can be tackled individually.

The CAROBS node architecture, redefined using tools of Queueing theory, contains three building blocks *SOA Switching Matrix* that ensures optical signal switching and two QS. For the sake of simplicity, these QS are define as the IQS and Output Queueing systems (OQS). The IQS tackles the contenting burst trains through a limited number of O/E blocks. The number of required O/E blocks depends on the IQS offered load. The offered load of the IQS can be quantified as $\alpha_{\text{BUF}} = \text{BP}.\alpha$, where α represents the total node offered load, α_{BUF} represents the offered load to the IQS, and BP represents the buffering probability. Using the α_{BUF} values, the number of O/E blocks can be calculated with respect to the SLE which means the streamline effect evaluation must be carried out for cases when flows are merging from less than five directions [44].

If there are flows merging from more than five directions, i.e., node degree is higher than 5 then the M/M/k/0 model can be applied [44]. Then the Erlang B formula can be used in order to obtain the BBP of IQS. The evaluation of the BBP enables us to obtain an approximation of the number of O/E blocks that are necessary in order to enable the loss-less mode. The OQS provides the burst train buffering and allows the traffic from client networks α_{AGG} to be aggregated. The behaviour of both QS are driven by the *SOA Manager*. Since the OQS receives burst trains in the form of an electrical signal, the OQS can store buffered burst trains in electrical memory. Current electrical memories provide only limited space, i.e., a limited number of burst trains can be stored; however, for the sake of simplicity and the CAROBS proposal [49] compliance, electrical memory is assumed to be unlimited in this analysis.

This simplification is caused by the modelling approach because real systems are not possible to built at the moment. Regards the modelling, every model relies on a number of approximations, the most crucial approximations in this analysis relate to the input traffic characteristics. The input traffic can be modelled using different input arrival processes and distributions of packet size. Some models are applicable only to a specific input traffic distribution or packet size distribution while some of them are more generic. In this analysis, the packets arriving from the connected network are assumed to follow a

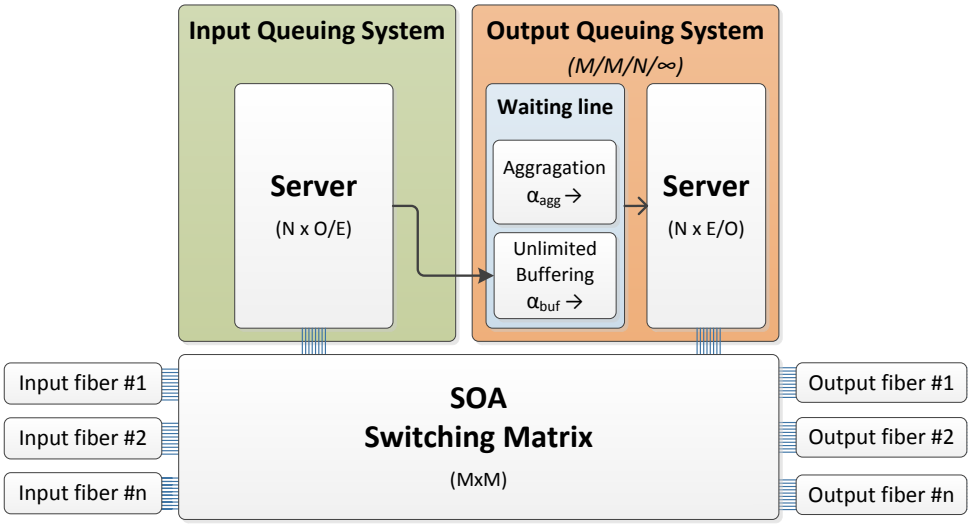


Figure 5.1: Simplified block structure of the CAROBS node architecture in terms of QS [33]. It contains two QS and one SOA Switching matrix. The IQS is responsible for the optical signal detection and sends the burst trains to the electrical buffer once they are detected O/E. The second QS controls the buffered trains and schedules them along the input traffic to the optical layer through a limited number of lasers E/O using LAUC-VF.

Poisson distribution at a rate of α packets per second so that the inter-arrival time between packets is a negative exponential distribution with parameter α. Depending on the car triggering levels [49], a burst train is created. CAROBS uses both triggering types (time, space)[49], therefore, car assembly tends toward Gaussian distribution asymptotically [40] according to the central limit theorem [59]. Using Kendall's notation [33], the OQS is classified as $M/M/N/\infty$ where N is the number of E/O blocks, see Figure 5.1.

5.1.2 CAROBS dimensioning formulation

The key characteristics of the CAROBS framework is its loss-less mode of operation, i.e., no burst must be dropped. In other words, there are always enough O/E blocks when contention occurs. On the other hand, O/E blocks are expensive from both an operational and installation perspective hence it is reasonable to minimize their number. Minimizing the number of O/E blocks implies that the effects of merging flows and secondary contention must be a priori minimized. As long as the secondary contention is the only effect of merging flows, it can be taken care by minimizing the merging flow effects. Therefore, next focus shifts towards to the classification of merging flows and their impact on burst buffering.

Currently, traffic in all-optical networks is distributed using both the RWA and grooming routing and wavelength assignment (GRWA) approaches [25], [60], [61]. In OCS networks, it is vital because there is no sub-wavelength scheduling; bandwidth sharing is achieved using traffic grooming in the electrical domain, so there is no contention in op-

tical domain. This means, the RWA is used in cases when optical domain needs to be partitioned, and the GRWA is used when finer usage of optical domain is necessary with use of electronic grooming. These two formulations are very helpful as a starting point for considering OBS RWA [26], [62]–[64]. The performance of OBS RWA algorithm is modelled using burst loss probability (BLP). The classical way to write the RWA using mathematical program is recalled in, e.g., [65], [66].

In order to write a basic GRWA for a given network and CAROBS framework, the network representation as graph is necessary thus a graph $G(\mathcal{V}, \mathcal{L})$ is defined. Let D be the traffic matrix which defines the amount of required bandwidth between any two nodes $s, d \in V$. Usually, the number of wavelengths is limited to $|\Lambda|$, assuming $\lambda \in \Lambda$. Then the objective function can be formulated as:

$$\min \sum_{(s,d)\in V^2 : s\neq d} \sum_{\substack{\ell\in L \\ \lambda\in\Lambda}} y_{\ell,\lambda}^{sd} + \theta \sum_{(s,d)\in V^2 : s\neq d} e_{sd} \tag{5.1}$$

where $y_{\ell,\lambda}^{sd}$ is a decision variable: it is equal to one if the required bandwidth flow ϕ from s to d is assigned on wavelength λ and link ℓ, and 0 otherwise. The e_{sd} is a variable representing how much of traffic could not be routed because there is not enough wavelengths in Λ to support all traffic D and θ is a objective function parameter. This objective function is subject to:
For all $v, s, d \in V$,

$$\sum_{\ell\in\omega^+(v)} \phi_\ell^{sd} - \sum_{\ell\in\omega^-(v)} \phi_\ell^{sd} = \begin{cases} D_{sd} - e_{sd} & \text{if } v = s \\ -D_{sd} + e_{sd} & \text{if } v = d \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

where $\omega^+(v)$ is the set of egress links of v and $\omega^-(v)$ is the set of ingress links of node v.

$$\sum_{s,d\in V} \phi_\ell^{sd} \leq y_{\ell,\lambda}^{sd} C, \qquad \lambda \in \Lambda, \ell \in L \tag{5.3}$$

$$\sum_{s,d\in V} \phi_\ell^{sd} \leq C, \qquad \ell \in L \tag{5.4}$$

$$\sum_{s,d\in V} y_{\ell,\lambda}^{sd} \leq 1 \qquad \lambda \in \Lambda, \ell \in L \tag{5.5}$$

$$\phi_\ell^{sd} \geq 0, e_{sd} \geq 0 \qquad \{s, d\} \in V, \ell \in L \tag{5.6}$$

$$y_{\ell,\lambda}^{sd} \in \{0, 1\} \qquad \{s, d\} \in V, \lambda \in \Lambda, \ell \in L. \tag{5.7}$$

where C is the wavelength capacity (assumed to be the same for all wavelengths).

The main drawback of this formulation is, this formulation does not take care of the merging flows and the sub-wavelength granularity of OBS. The number of merging flows should be minimized as much as possible in order to maximize the occupancy of the simple, already merged flows. Thankfully, it can be done in the online mode using load balancing algorithms maximizing the SLE [44]. However, such an algorithm corresponds to a heuristic, hence, it does not provide a globally minimal solution from the perspective of the number of O/E blocks. Additionally, the number of O/E blocks depends on the characteristic of the node offered load, implication of M/M/N/0 model [33]. The stream of incoming burst trains can be classified as a process with random inter-arrival intervals [40], therefore, it may result in infinite waiting in the electronic memory for a certain level of offered load. The infinite waiting is caused by the aforementioned premise of infinite memory, however, in real networks it would result in a burst loss because of limited electronic memory. Consequently, for the minimization of the number of O/E blocks, the offered node load must be engineered to avoid excessive buffering.

The conclusion of this section is as follows. Both QS must be dimensioned properly, otherwise burst loss can be experienced. Both QS dimensioning relates to the maximal CAROBS node load which can be formulated as the input traffic intensity ρ, so that $\rho \equiv \alpha/\mu$. The stability condition for the CAROBS node is then stated as $\rho < 1$, which can be written as $\alpha < \mu$ [33]. Here, α stands for total node offered load and μ represents the node intensity of service, i.e., how much traffic a node can transmit. The node offered load α is equal to the sum of offered loads from each tributary flow. It is worth noting that the stability condition applies to the system with merging flows, otherwise applies the SLE. For the SLE, the stability condition changes to $\rho \leq 1$.

Now one may wonder what is the optimal value of ρ for various numbers of merging flows because then constraints (5.3) can be reformulated to:

$$\sum_{(s,d)\in V^2} \phi_\ell^{sd} \leq y_{\ell,\lambda}^{sd}.C.K \qquad \lambda \in \Lambda, \ell \in L_{\text{selected}}, \tag{5.8}$$

where L_{selected} represents the number of outgoing links of the merging node V, e.g., ℓ_0 in Figure 5.2, which concentrates traffic from a number of merged flows and K is the coefficient of stationary threshold that ensures that the egress link is used efficiently $C.K \rightarrow \rho_{\max}$, i.e., no excessive buffering nor burst loss occurs.

Constraints (5.8) imply that the maximum node load is limited. Then the stability condition is valid, additionally the electronic memory is not overloaded.

5.2 Simulations

Following the discussion on the optimization of the number of O/E blocks, BP cannot be estimated using either the M/M/k/k model, or the SLE because both provide information about BBP but not about BP. There is a clear relation between BP and BBP but the main difference is that BP is influenced by secondary contention. The BP value can be quantified either mathematically or empirically using simulations. In this analysis, an approximation model relying on simulations, prior to the design of a mathematical model in the future, is discussed. The main focus is on the basic node behaviour under various conditions, using the topology described in Figure 5.2. The most important node, the node under study, is marked as merging node v, see Figure 5.2. There is also destination node d where all traffic flows, from sources $s_\bullet$, are destined. The number of sources changes, so four scenarios with two to five merging flows are evaluated. The maximum of five merging flows was chosen because the maximum node degree that is considered is six, i.e., five merging flows in [67]. Simulations were carried out using OMNeT++ simulator and CAROBS models [53]. Source nodes $s_\bullet$ were supplied with traffic generated according to the Poisson distribution. The generated payload packets of constant size (100 kb) defining the flow were supplied to aggregation queues to generate bursts. It is assumed that electrical storage capacity is unlimited. JET [13] was used as a signalling protocol and LAUC-VF algorithm [68] for burst assembly.

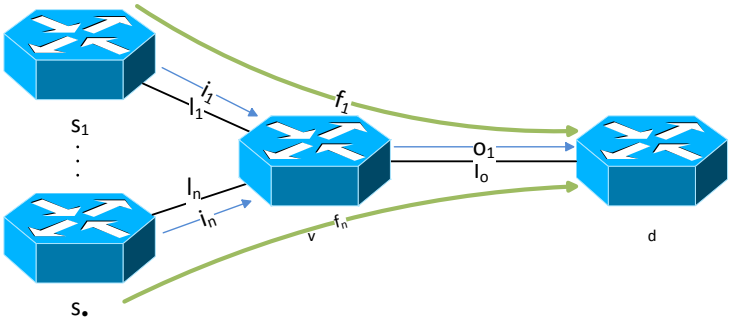


Figure 5.2: The elementary topology used for one node behaviour evaluation. The number of sources was changed as is depicted here by $s_1, s_2, \ldots, s_\bullet$.

Traffic analysis is quite comprehensive if one wants to reach a specific level of accuracy. The number of simulations must be as high as possible to obtain statistically accurate results. Accuracy of results is obtained through the repetition of identical simulations and by changing the input load patterns. However, exact results cannot be obtained in real time, therefore, a limited number of simulations were used.

First, the number of simulations in order to obtain results with statistical accuracy are addressed. The upper number of identical simulations is bounded by the time that can be invested into the simulations. In these simulations, 25 identical simulations are carried

out with different patterns of node offered load. The main task of the simulations is to verify the impact of the number of merging flows (MF) and node load α on the buffering probability BP(α, MF) and buffering delay BD(α, MF) of a buffered burst train. The node load α is equal to the sum of loads offered by each tributary flow. Then, the average maximal offered load provided by each source is 1/(MF+1) erl. The term average maximal offered load represents the average value of offered load among the identical simulations for one simulation at the given node offered load α. With regard to this approach, there must be enough data for the proper statistical evaluation, hence, valid interpretations.

In addition to the number of merging flows and the node load, an evaluation for wavelength data rate 1, 10 and 40 Gbps is performed. The Section 5.3.2 tackles the O/E block sharing in WDM networks to verify the impact of the wavelength number $|\Lambda|$ on BP(α, MF) $\rightarrow$ BP(α, MF, $|\Lambda|$) and BD(α, MF) $\rightarrow$ BD(α, MF, $|\Lambda|$). In this experiment the number of wavelengths $|\Lambda|$ is changed from 1 up to 60.

5.3 Results

This section goes down to the results obtained from OMNeT++ simulations. In the begging, Section 5.3.1, simulations are restricted to only one wavelength and wavelength data rates are changed. Therein, first the focus is on the stability of measured parameters and their confidence. Based on the results, recommendation for the number of O/E blocks in order to ensure loss-free mode of operation is formulated. These results also enable CAPEX studies to be carried out. Also these results open question about viability of one wavelength systems, i.e. wavelength sensitive O/E blocks.

Therefore, in Section 5.3.2, the multiple wavelengths (WDM mode of operation) and colour-less O/E blocks are simulated. These results favors the WDM in CAROBS networks because WDM mode decreases number of O/E blocks that are necessary for the same load as for one wavelength scenario significantly. Additionally these results open a very promising deployment scenario that is unique for CAROBS.

5.3.1 One wavelength evaluation

One-wavelength transmission systems were common before the emergence of WDM. This section returns to the one-wavelength system because of its simplicity, i.e., less degree of freedom in the analysis. Less freedom means a smaller number of parameters that can change. Such a (lower) number of parameters is vital for the stability discussion of results as other parameters are fixed. The one-wavelength analysis consists of two main steps. In the first step, the number of identical simulations is fixed, so the following results are a

trade-off between accuracy and the overall time of running the simulations. In the second step, the predefined simulations are repeated n times then results are evaluated.

In order to find the trade-off between accuracy and overall simulation running time, simulations for three different node loads $\alpha = \{0.1, 0.5, 0.9\}$ erl are conducted. For each of these three-node loads, 2,000 simulations with different offered load patterns are performed. Additionally, this analysis is conduct for data rates 1, 10 and 40 Gbps. In this analysis, only the stationary simulations are used and they are tested using the mathematical tool of hypothesis testing. For formulate a claim that the signal is stationary when the values of BD and BP do not depend on time, i.e., statistical properties do not change in time during the simulation. This test is applied to the whole simulation time interval, in addition to the sub-intervals. Each of the sub-intervals contain 120 samples. In the test, it is assumed that the values of BD and BP in time can be approximated by a linear regression line, which is quantified using time vector X, the coefficients of linear regression b_0, b_1 and the coefficient of linear regression error e_i which must be minimal. Then, a null hypothesis $H0$ claiming that the signal is stationary if every sub-interval is stationary can be formulated. Sub-interval stationarity is characterized by coefficient $b_1 = 0$. If $b_1 \neq 0$, then such sub-intervals are not stationary, so the whole time interval, i.e., the simulation cannot be used for further stability analysis. The core of the linear regression analysis is shown in (5.9)-(5.12).

$$Y_i = b_0 + b_1 x_i + e_i \qquad i = 1, 2, \ldots, n \tag{5.9}$$

$$b_1 = \frac{n\sum_{i=1}^{n} x_i^2 Y_i - \sum_{i=1}^{n} x_i \sum Y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \qquad b_0 = \frac{\sum_{i=1}^{n} Y_i - b_1 \sum_{i=1}^{n} x_i}{n} \tag{5.10}$$

$$s^2 = \frac{\sum_{i=1}^{n} Y_i^2 - b_0 \sum_{i=1}^{n} Y_i - b_1 \sum_{i=1}^{n} x_i Y_i}{n-2} \tag{5.11}$$

$$\frac{|b_1|\sqrt{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}}{s} \geq t_{n-2}(\Psi) \tag{5.12}$$

where n is the number of verified samples in the sub-interval and the number of sub-intervals for the hypothesis testing of the whole simulation interval. Then, the decision on stationarity is valid with a level of confidence Ψ. If the critical value (5.12) is higher than the coefficient of Student's distribution, the hypothesis $H0$ does not apply, i.e., BD or BP is not stationary and the simulation cannot be used for further evaluations. The valid

set of simulations is then used for the evaluation of the number of identical simulations. This approach is also used in the next analysis of the maximal node load.

In order to find the minimal number of identical simulations, 2,000 simulations are performed. The statistical properties of datasets representing various numbers of identical simulations can be seen in Figure 5.3. The statistical properties do not change after 1,000 identical simulations, therefore, for the sake of readability, only limited number of repetitions are depict in Fig.5.3. One can see that in most cases after 25 identical simulations, the mean value and variance do not change significantly against their values with 1,000-simulation case. Therefore, 25 identical simulations are going to be used for all simulations in this paper. As long as there are only 25 identical simulations, the results can be corrected using the coefficients of Student's distribution.

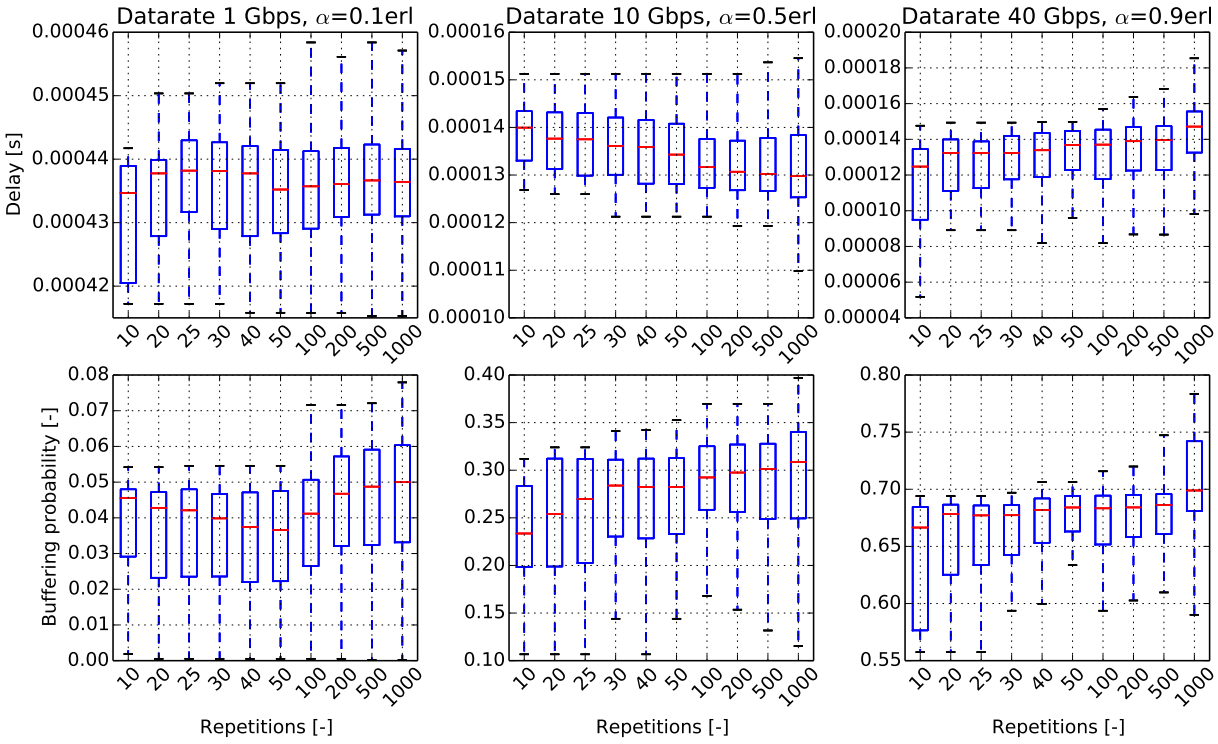


Figure 5.3: Statistical properties of BD and BP captured using boxplots for various node loads and various data rates. In the upper row, characteristics of buffering delay are depicted. In the bottom row, characteristics of buffering probability for a different number of identical simulations are shown.

The impact of the number of merging flows to the BD and BP is an extremely important aspect of the analysis which is carried out using the dataset containing 2,000 simulations. For the sake of simplicity, only the case for the node load 0.5 erl is depicted, however, all other simulation schemes led to the same conclusion that BD does not depend on the number of merging flows, see Figure 5.4, only on the node load, see Figure 5.6a). This conclusion comes from the Poisson character of the merged flows. Therefore, in the

delay analysis, $\mathrm{BD}(\alpha, \mathrm{MF}) \equiv \mathrm{BD}(\alpha)$ can be assumed.

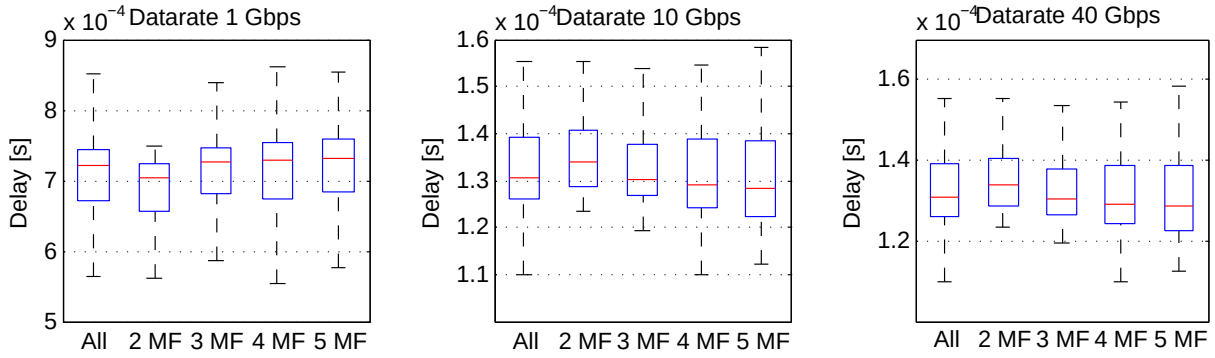


Figure 5.4: Buffering delay test of dependance for node load $\alpha = 0.5$ erl showing Buffering delay does not depend on the number of merging flows. The routing policy does not avoid scenarios with a higher number of merging flows for the same level of offered load when the end-to-end delay is the main concern.

The same dataset of 2,000 different patterns was used for the BP analysis with the results for node load 0.5 erl depicted in Figure 5.5. Other node load results led to the same conclusion. According to the results depicted in Figure 5.5, it can be seen that above 2 MF the difference of BP is negligible; therefore, BP can only be defined for two or more MF as two different parameters in the next analysis. This is indirect contradiction to the M/M/k/k with an inclination to the SLE for small MF scenarios.

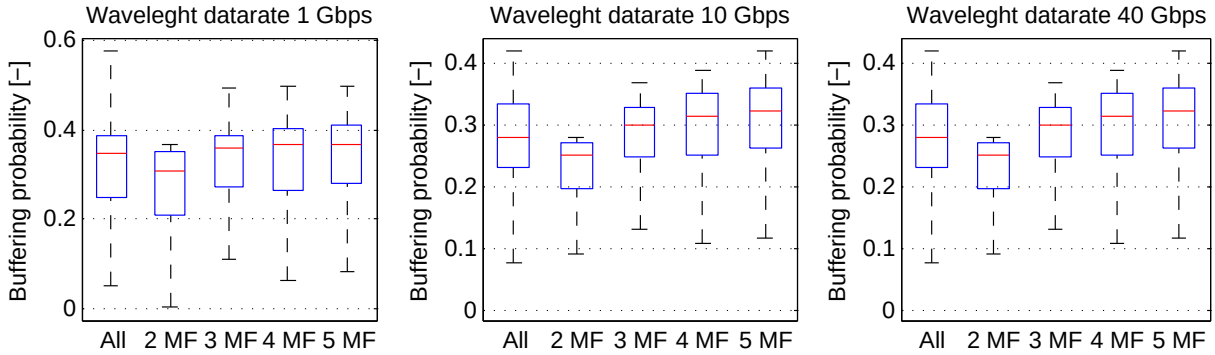


Figure 5.5: Buffering probability dependance test on the number of merging flows. The test was carried out for the node load $\alpha = 0.5$ erl. This test shows that the Buffering probability does not depend on the number of merging flows for $MF \geq 3$.

In the performance analysis of the CAROBS system using one wavelength, 25 schemes were performed with a variety of offered load patterns for load $\alpha = [0.1; 1.02]$ erl with equidistant step 0.02 erl for $\alpha = [0.1; 0.95]$, and with step 0.01erl for $\alpha = [0.95; 1.02]$ erl. The scheme using 2-5 MF and three data rates 1, 10 and 40 Gbps of a wavelength was retained and BD and BP stationary tests were evaluated to obtain the stationary threshold

K. The values determining the stationary thresholds are captured in Table 5.1. Therein, a column F_{BW} is added to expresses the stationary threshold in terms of the bandwidth that must be free to avoid excessive buffering for static traffic.

Table 5.1: Table of load thresholds and spare bandwidth. F_{BW} represents the spare bandwidth, which cannot be used to keep the node stable in a long-term perspective.

	1 Gbps		10 Gbps		40 Gbps	
MF	K	F_{BW}	K	F_{BW}	K	F_{BW}
2	0.89	110 Mbps	0.93	700 Mbps	0.96	1.6 Gbps
3	0.88	120 Mbps	0.89	1.1 Gbps	0.89	4.4 Gbps
4	0.88	120 Mbps	0.89	1.1 Gbps	0.89	4.4 Gbps
5	0.87	130 Mbps	0.88	1.2 Gbps	0.88	4.8 Gbps

The values of BD(α) and BP(α, MF) are seen in Figure 5.6. The differences among BD(α) for various wavelength data rates are shown in Figure 5.6a). When the end-to-end delay is the routing concern, it can be seen that the node load cannot exceed ≈ 0.9 erl, otherwise BD exponentially increases. Further, the value of the load ≈ 0.9 erl is equal to the stationary threshold which means that the node could not ensure the loss-less mode permanently. It would eventually lead to burst loss. The second monitored parameter BP(α, MF) is seen in Figure 5.6b) where there are depict only two and five MF cases. The notably high BP is evident even for the low node load from this figure. Such a situation is not vital for production networks; therefore, wavelength dependent O/E blocks are not a viable way how to deploy CAROBS in WDM networks. On the other hand, this situation provides very useful data that can be used for further analysis as an upper bound when the number of O/E blocks is the main concern.

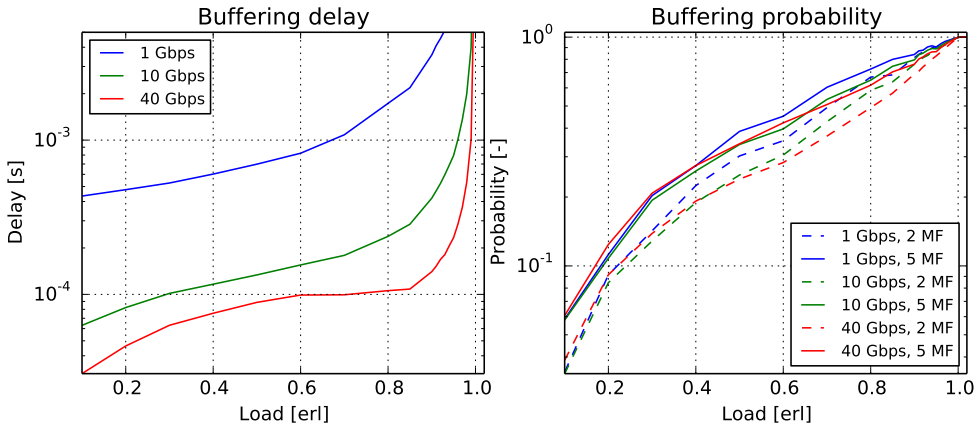


Figure 5.6: The BD and BP rely on the value of the offered load. Below 0.9 erl the BD is in scale of μs, however, above 0.9 erl, it significantly increases. The BP is more proportional to the value of the offered load α. Therefore, when engineering the number of regenerators, BP is the main objective.

The accuracy and results of BP are crucial when estimating O/E blocks. Each O/E

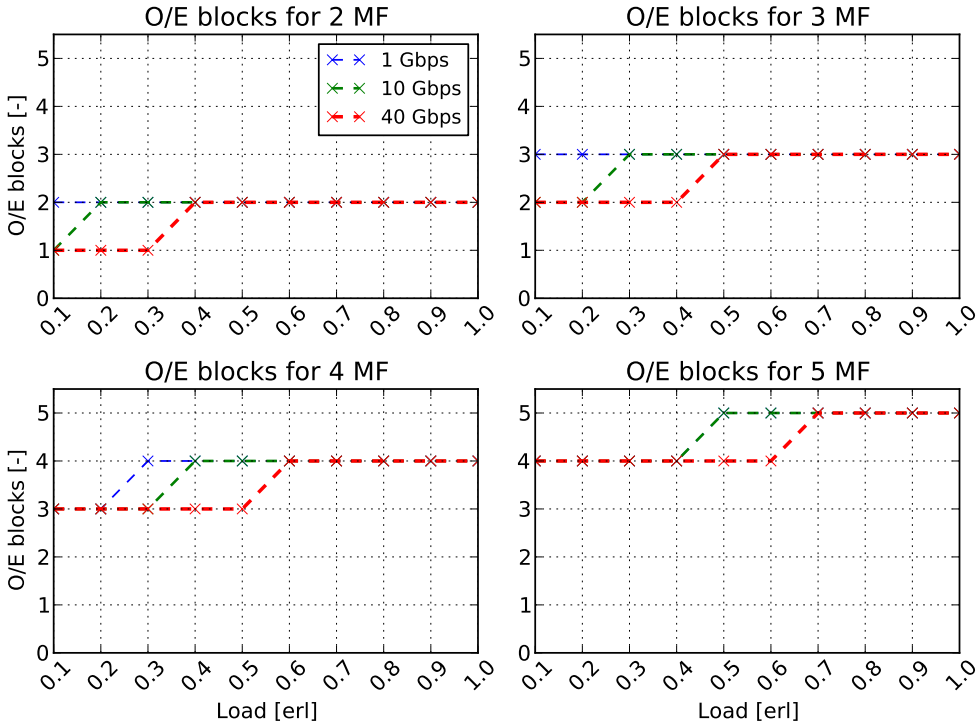


Figure 5.7: Dependance of the number of O/E blocks for various MF scenarios and wavelength data rates. The most resilient case is for 2 MF. This scenario is also the cheapest, though sometimes at the expense of wavelength greediness. The evaluation of $R(\alpha, MF)$ is calculated from Figure 5.6b).

block can be used by only one burst at a time, and only one burst train can come, at a given time period, each wavelength only carries one single burst train. At this moment, the number of buffered bursts depend on the offered load α and the number of MF, BP(α, MF). The BP(α, MF) specifies the value of the probability that a new incoming burst train will be blocked by another burst train (burst contention); in the worst-case scenario, it can be blocked by a rescheduled burst train (secondary contention). The worst-case scenario results in the corner case of equal numbers of O/E blocks and MF. O/E block measurement is depicted in Figure 5.7. One can be see that it is necessary to install at least $MF - 1$ O/E blocks even for a low load. These graphs quantify the O/E block requirements, so these graphs can be used in further studies of CAROBS GRWA. As with the BP, the O/E blocks graphs reveal a high demand on the number of O/E blocks, even for a low load; therefore, it is not viable for real CAROBS deployment in WDM networks where deployment requires a specific number of O/E blocks per wavelength depending on the α at the wavelength.

The total number of O/E blocks required for the CAROBS network and given traffic D can be calculated through the CAROBS GRWA extension which gives information about virtual routing (different routing for different wavelengths). Each wavelength can have a different number of merging flows (MF_λ) and a different node offered load α_λ. The

evaluation of the number of O/E blocks can be formulated as follows:

$$\sum_{\substack{v\in V\\ \lambda\in\Lambda}} R(\alpha_{v,\lambda}, \mathrm{MF}_{v,\lambda}) \tag{5.13}$$

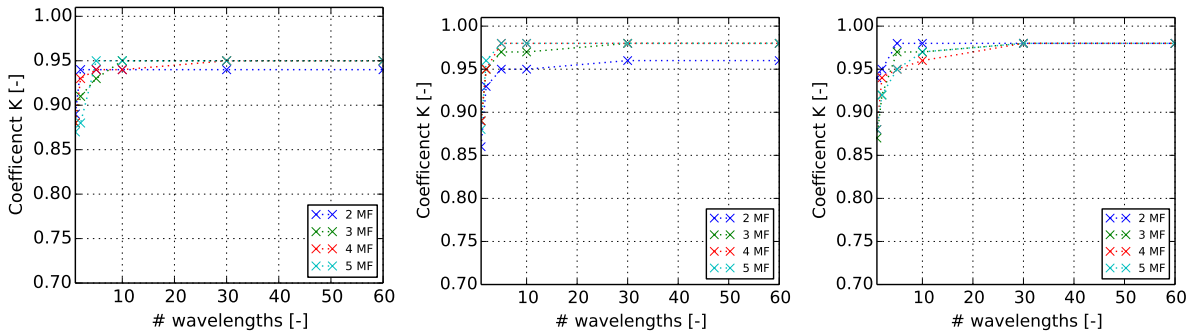
where function R represents the graphs depicted in Figure 5.7, i.e. the requirement on the O/E block to deliver the loss-free mode of operation. Such a formulation can be used for CAPEX or OPEX studies where it can partially define the objective function in order to minimize monetary sources while maintaining an appropriate quality of transmission.

5.3.2 Multiwavelength evaluation

Following the results obtained in the previous section, the focus of this section is only on the BD, BP and the number of shared O/E blocks. Additionally, various schemes of the number of wavelengths in the CAROBS WDM is considered. First, the stationary threshold of the CAROBS WDM system with a various number of wavelengths $|\Lambda|$ will be evaluated, followed by the measurement of colour-less O/E blocks. The term "Load" will be used in all figures, however, the meaning is modified compared to the previous section where it meant the total utilization of one wavelength. From now on, the term "Load" represents the overall utilization of all wavelengths in a specific link – link utilization. For example, a link supporting 10 wavelengths, where only one wavelength is utilized by a 1 erl traffic, according to the notation used in the previous section, means that Load is equal to 0.1 erl. K is redefined the same way as Load.

The stationary threshold of CAROBS WDM was evaluated based on the dataset of the CAROBS WDM system where the $|\Lambda|$ was varied, and the number of merging flows MF and node offered load α was changed. The accuracy of each step, as defined in Section 5.3.1, is evaluated using the mean value analysis (MVA) approach. The MVA was carried out so the final value is uncertain with less than 5% of probability. The dependence of the coefficient $K(|\Lambda|)$ is illustrated in Figure 5.8.

The inclination of coefficient $K(|\Lambda|)$ to the value one is seen, however, in the studied range of wavelengths the coefficient $K(|\Lambda|)$ does not meet it. It results into the gap of bandwidth F_{BW} that cannot be used for static traffic but could be used for frequently bursting short term flows which cannot result in excessive buffering. The values of coefficient $K(|\Lambda|)$ delimit a working area where the CAROBS WDM system can be provisioned. The BD and BP aligned to the working area are investigated next. Their graphs are depicted in Figure 5.9. BP is captured for two and five MF in Figure 5.9abc), BD$(\alpha, |\Lambda|)$ is depicted in Figure 5.9def) without respect to the number of MF, because of the Poisson character of merging flows, see Section 5.3.1. Both BP and BD improved significantly as



(a) Wavelength datarate 1 Gbps. (b) Wavelength datarate 10 Gbps. (c) Wavelength datarate 40 Gbps.

Figure 5.8: Dependance of reduction coefficient K on the number of wavelengths. The value of coefficient $K_{bps}(\lambda, |\Lambda|) \to 1$ is not equal to one, i.e., there is some free capacity necessary to keep the system stationary.

the number of wavelengths $|\Lambda|$ increased. It is an excellent indicator that further research on O/E block sharing is the right direction.

The gap between the two and five MF scenarios is worth studying as it takes on importance as the wavelength data rate increases. The gap is the direct result of SLE, i.e., suppressed secondary contention and it can be used to minimize the number of deployed O/E blocks in the network. The objective can be formulated as the minimization of MF and maximization of the stream lining. In this section, it is obvious that SLE is highly important for the CAROBS node and its performance.

The number of O/E relates to BP as in the previous section. The graphs are depicted in Figure 5.10. The one-wavelength scenario is captured here to depict the upper bound described in the previous section. One can see, that O/E blocks sharing among wavelengths can significantly reduce their necessity for the same level of offered load. It is significant to note that it is not necessary to install any O/E blocks for buffering up to a specific level of node offered load. Such a threshold can be used for designing simple CAROBS nodes with minimal requirements. On the other hand, it allows to design CAROBS nodes which tackle most contention.

The shared O/E blocks approach allows us exploring the area of CAROBS nodes deployment with no O/E blocks. It is promising for networks with centralized buffering as it allows traffic routing where it is not necessary to install any O/E blocks for contention resolution at a particular node. Geographically extensive deployments, where some nodes can be designed for a low load, can be deployed more cheaply and easily. That notwithstanding, more powerful nodes can be installed in data centers allowing contention resolution through O/E blocks. It is important in future studies to return to the regeneration of optical signal because of optical impairments and these results give us a good starting point. This section showed that the CAROBS WDM with shared O/E

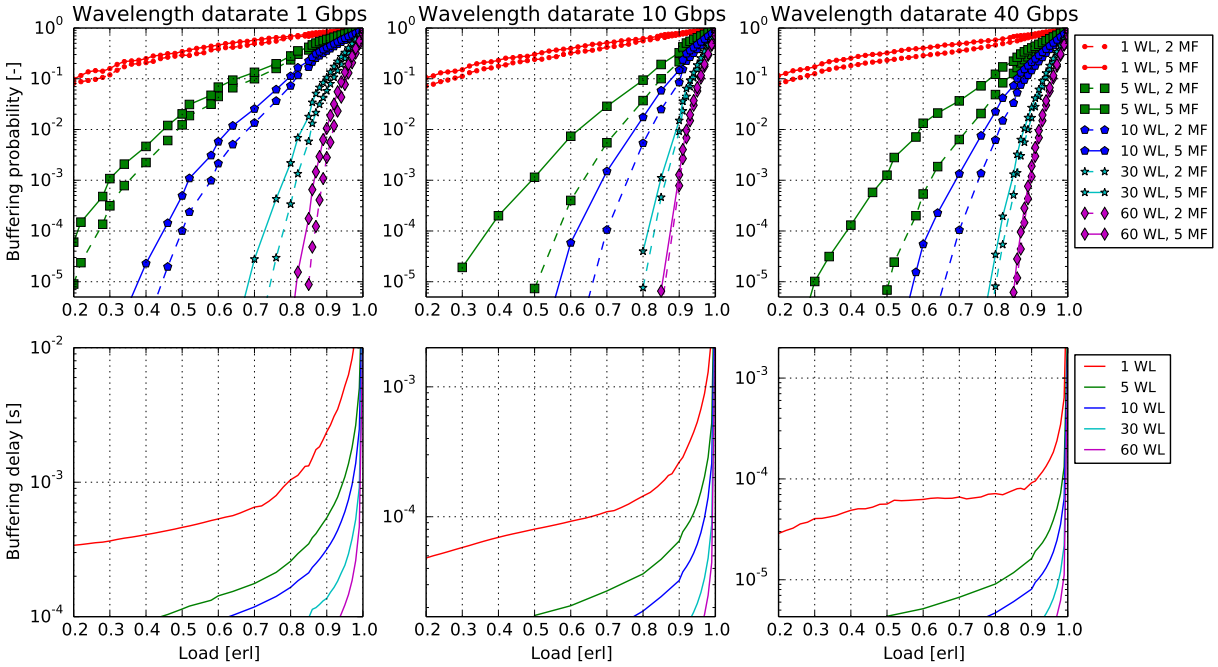


Figure 5.9: A comparison of CAROBS using the WDM system. The upper set of figures depict the buffering probability for six different wavelength sets and two and five MF. The lower set of figures captures the average buffering delay with no respect to the number of MF. Both sets offer clear evidence for shared O/E blocks deployment.

blocks has minimal requirements on O/E blocks for contention resolution. Therefore, it is worth investing more research into optimal routing and O/E block installation in further work.

5.4 Implications to the design of CAROBS GRWA algorithm

The CAROBS GRWA gives information about traffic routing and necessity of resources (the number of O/E blocks, etc.) installed at nodes in a network. The analysis in the previous section shows that the colour-less detectors are mandatory in order to reach low level of CAPEX and OPEX due to lower number of O/E blocks which are installed. Additionally, the system using colour-less detectors allows to route traffic in a network in such a way that in some situations, node load is small and the number of wavelength is high, contention can be fully mitigated.

In order to find such a routing an GRWA algorithm must be properly formulated. The main focus is the CAPEX and OPEX minimization hence the installation of O/E blocks must be minimized. The CAPEX and OPEX minimization is tightly bounded to results captured in Figure 5.10. These results can be used as sum of all O/E blocks across nodes

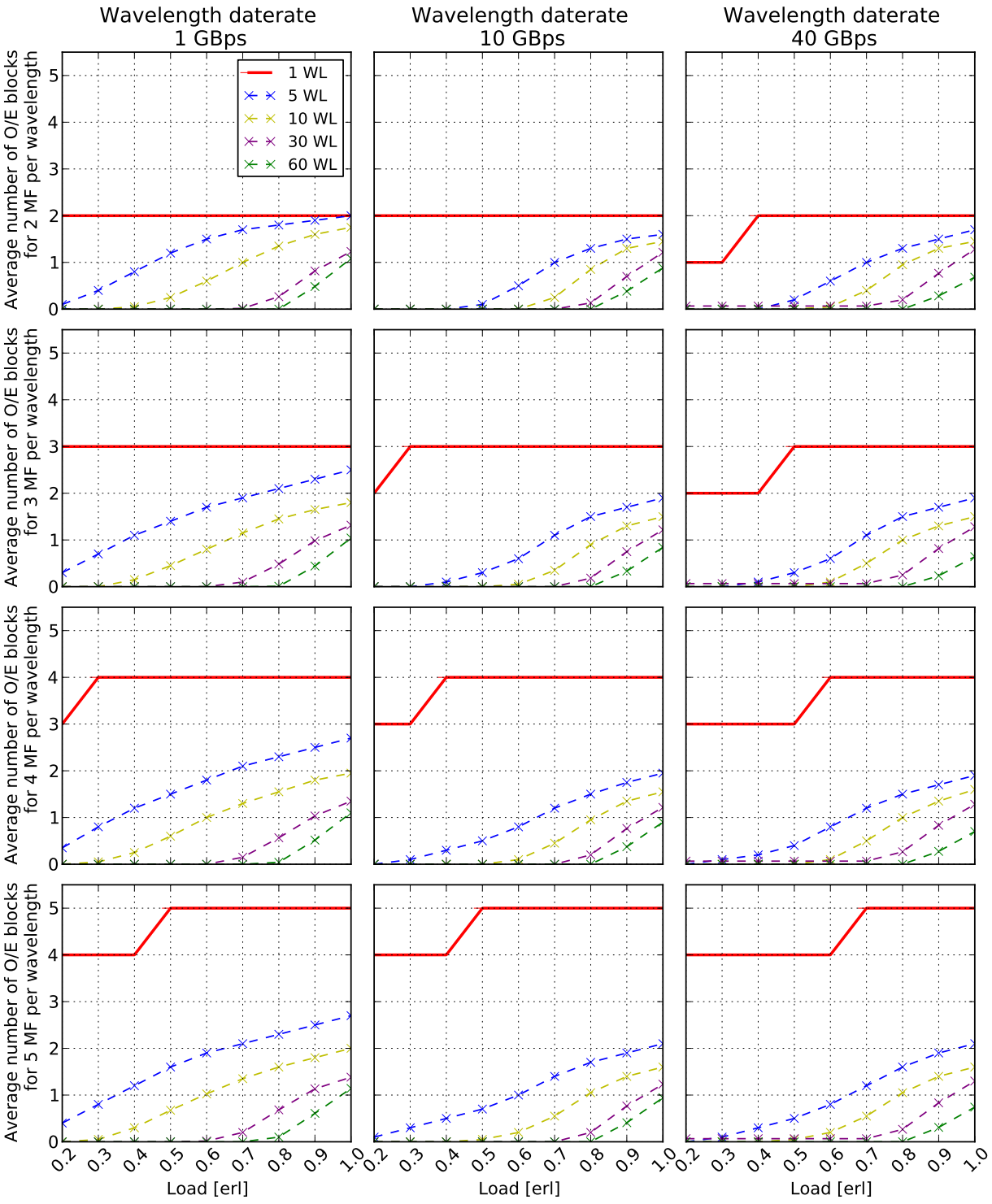
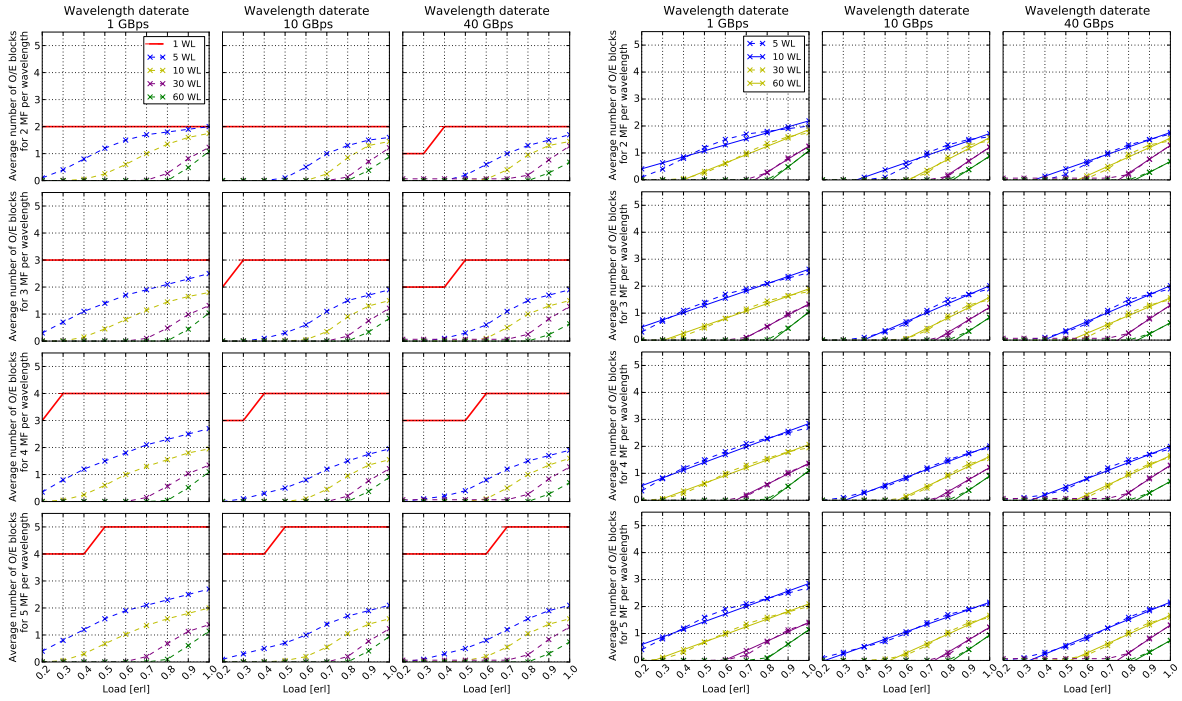


Figure 5.10: Estimation of the average number of O/E blocks per wavelength with respect to the egress port load. The higher the number of wavelengths shared by one O/E block, the lower the total number of O/E blocks. The upper boundary for the system containing a higher number of wavelengths is lowered because it is over the threshold K, i.e., a load that should not be reached.

in a network. However, these results are very hard to use because they have non-linear nature. It means with small change of the Load they do not change linearly which is necessary for linear programming formulation that is selected for optimizations in this thesis. Therefore, the linearisation is necessary, and the linearised results are depicted in Figure 5.11(b). The price for linearisation is in the accuracy of results hence the non-linear lines were approximated with a line that gives minimum error.



(a) Original function of average number of O/E blocks

(b) Linearized function of the average number of O/E blocks

When the parameter describing dependence of the number of O/E blocks on the Load is linearised, the number of O/E blocks can be calculated using the (5.14) for $x > b_{thresh}$. The b_{thresh} represents the node load when some contentions can occur, so O/E blocks are necessary.

$$\bar{O/E} = b_0 + b_1 x \tag{5.14}$$

At this point extra attention must be paid because the linearised data are normalized to the number of wavelengths $|\Lambda|$ in the WDM transmission system. Therefore, the accurate number of O/E blocks is function of Load, the number of wavelengths $|\Lambda|$, and most importantly to the number of merging flows MF. The total number of O/E blocks per one node can be enumerated as (5.15).

$$O/E(\Lambda, \alpha, MF) = |\Lambda|.\left(b_{0,|\Lambda|,MF} + b_{1,|\Lambda|,MF}\alpha\right) \tag{5.15}$$

The dependence of the number of O/E blocks is caused by the dependence of parameters b_0, b_1 to the number of wavelengths and merging flows. This dependence of b_x parameters is captured in the Figure 5.4

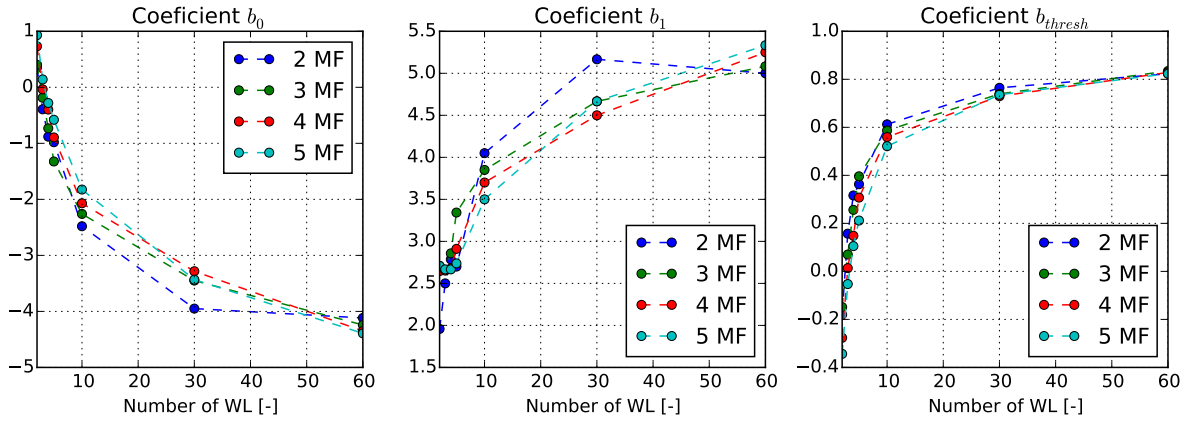


Figure 5.11: Values of b parameters with respect to the number of wavelengths, at 10 Gbps per wavelength.

5.5 CAPEX and OPEX minimization algorithm for CAROBS WDM networks

The CAPEX can change in time with advances of technology, but this change is very hard to predict. A better way how to optimize CAPEX is to minimise the use of expensive function blocks. In case of CAROBS WDM, this means minimization of blocks delivering loss-less mode of operation, i.e., O/E blocks, because the CAROBS architecture does not contain any extra blocks compared to the original OBS architecture [13]. Due to the minimization of the number of the O/E blocks the OPEX is going to be minimized as well. O/E blocks are responsible for the conversion of the contenting optical burst into electrical domain hence O/E blocks contain electronics that consume electrical energy. This consumed electrical energy increases with the increase of the optical bandwidth[69].

Based on the information in the previous section, the integer linear program (ILP) formulation minimizing CAPEX and OPEX can be written. This formulation is going to be based on the Maximal throughput approach, see Section 3.2.2, using multi-objective objective function which is going to deliver the O/E blocks minimization and maximization of the supported traffic.

5.5.1 Objective function

Objective function determines the main scope of optimization algorithm, so the (5.16) contains two parts which are balanced using penalties PENAL_{xxx}. The preference of the solution can be formulated towards to the either unsupported traffic PENAL_{D} or network cost PENAL_{OE}. Eitherway, each solution of the ILP program highly depends on the requested traffic D and physical properties of the optical network, such as Λ and network topology which can be represented as a graph $G(\mathcal{V}, \mathcal{L})$. The objective function is formulated as follows:

$$\min \left(\text{PENAL}_{\text{D}} \sum_{(s,d)\in V^2 : s\neq d} e_{sd} + \text{PENAL}_{\text{OE}} \sum_{v\in V} OE_v \right) \tag{5.16}$$

where e_{sd} is a variable representing how much of traffic could not be routed because there is not enough wavelengths in Λ in order to support all traffic D.

5.5.2 Constraints: Flow-control

The crucial set of constrains is based on the Maximization throughput problem formulation, see Section 3.2.2 that is extended with respect to the WDM. Therefore, the traffic RWA is captured by constrains:

$$\sum_{\ell\in\omega^+(v)} \phi^{sd}_{\ell,\lambda} - \sum_{\ell\in\omega^-(v)} \phi^{sd}_{\ell,\lambda} = \begin{cases} b^{sd}_{\lambda} & \text{if } v = s \\ -b^{sd}_{\lambda} & \text{if } v = d \\ 0 & \text{otherwise} \end{cases} \qquad v, s, d \in \mathcal{V}, \lambda \in \Lambda \tag{5.17}$$

where $\omega^+(v)$ is the set of egress links of v and $\omega^-(v)$ is the set of ingress links of node v.

$$\sum_{\lambda\in\Lambda} b^{sd}_{\lambda} = D_{sd} - e_{sd} \qquad s, d \in \mathcal{V} \tag{5.18}$$

$$\sum_{\lambda\in\Lambda} \sum_{s,d\in\mathcal{V}} \phi^{sd}_{\ell_o,\lambda} \leq |\Lambda|.C.K, \qquad \ell_o \in \omega^+(v), v \in \mathcal{V} \tag{5.19}$$

$$\sum_{s,d\in\mathcal{V}} \phi^{sd}_{\ell,\lambda} \leq C, \qquad \ell \in \mathcal{L}, \lambda \in \Lambda \tag{5.20}$$

$$\phi^{sd}_{\ell,\lambda} \leq y^{sd}_{\ell,\lambda} D_{sd}, \qquad \ell \in \mathcal{L}, \lambda \in \Lambda, s, d \in \mathcal{V} \tag{5.21}$$

where the $\phi^{sd}_{\ell,\lambda}$ is a non-negative variable representing the amount of traffic that is routed $s \to d$ at a specific wavelength λ, and the variable $y^{sd}_{\ell,\lambda}$ is the binary representation of

variable $\phi^{sd}_{\ell,\lambda}$. The variable $y^{sd}_{\ell,\lambda}$ is one when $\phi^{sd}_{\ell,\lambda}$ is positive. The constrains (5.19) represent the maximal node load when the node is still stable, and the coefficient K is based on (5.8). The constraints (5.18) are closely related to (5.17), and these constraints ensure maximization of routed traffic for each $s \to d$ request. The (5.20) restrict the maximal usage of each wavelength, and constraints (5.21) create the relation between variables $\phi^{sd}_{\ell,\lambda}$ and $y^{sd}_{\ell,\lambda}$. These variables are very important in order to mitigate minor routing problems related to loops, but for the sake of simplicity these constraints are not listed here.

5.5.3 Constraints: O/E counting

The Figure 5.11(b) and (5.15) give information how to enumerate the number of O/E blocks based on the node load, number of wavelengths $|\Lambda|_\ell$, and the number of merging flows MF_{ℓ_o}. In the previous section, the traffic routing is carried out, so an traffic routing framework is ready, and O/E blocks constraints are necessary in order to influence the flow-control.

Optical cross connect configuration

The configuration of the elementary MX reflects the RWA, i.e., how flows are routed. Each flow relates to a node v by variable $\chi^{sd}_{v,\ell_i,\ell_o}$ which is one when there is a sd flow at node v, connected from input port ℓ_i to output port ℓ_o, and at wavelength λ. Because the CAROBS enables traffic grooming, $\chi^{sd}_{v,\ell_i,\ell_o}$ variables can not be used for merging flows counting directly.

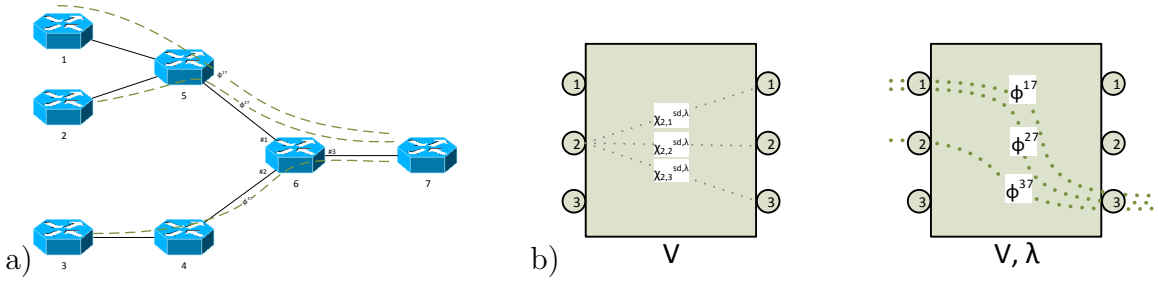


Figure 5.12: Visualisation of merging problem a) is the macro view and b) is the inside look at node **6**. Here, three flows are pictured all in the same colour (to underpin they are on the same wavelength) labelled by ϕ^{sd}. Most importantly, these flows correspond to the decision variables $\chi^{sd}_{v,\ell_i,\ell_o}$. This variable is one if such a connection is used by a flow ϕ^{sd} and zero otherwise. In the b), the decision variable $\chi^{sd}_{v,\ell_i,\ell_o}$ is depicted only for input port 4 in order to increase the figure readability.

The values of variables $\chi^{sd}_{v,\ell_i,\ell_o}$ are derived from the decision variable $y^{sd}_{\ell,\lambda}$ using the following formulation:

$$\chi^{sd,\lambda}_{\ell_i,\ell_o} = y^{sd}_{\ell_i,\lambda} y^{sd}_{\ell_o,\lambda} \qquad v \in V \setminus \{s,d\}, \ell_i \in \omega^-(v), \ell_o \in \omega^+(v) \tag{5.22}$$

which puts into relation the input and output port at specific wavelength. The drawback of these constrains is, these constraints are non-linear hence they are not possible to use in ILP and must be linearised. The linearised formulation is as follows:

$$\chi^{sd,\lambda}_{\ell_i,\ell_o} \leq y^{sd}_{\ell_i,\lambda} \qquad v \in V \setminus \{s,d\}, \ell_i \in \omega^-(v), \ell_o \in \omega^+(v) \tag{5.23}$$

$$\chi^{sd,\lambda}_{\ell_i,\ell_o} \leq y^{sd}_{\ell_o,\lambda} \qquad v \in V \setminus \{s,d\}, \ell_i \in \omega^-(v), \ell_o \in \omega^+(v) \tag{5.24}$$

$$\chi^{sd,\lambda}_{\ell_i,\ell_o} \geq y^{sd}_{\ell_i,\lambda} + y^{sd}_{\ell_o,\lambda} - 1 \qquad v \in V \setminus \{s,d\} \tag{5.25}$$

Merging flows calculation

The number of merging flows implicates the number of O/E blocks for every particular output port ℓ_o. The CAROBS uses sub-wavelength scheduling, so the variables $\chi^{sd}_{v,\ell_i,\ell_o}$ are problematic. For example, the situation in Figure 5.12 would result into 2 O/E blocks because there are two flows entering the node through port 1 and one flow entering node through port 2, and all flows are leaving the node by port 3. However, in this particular case only one O/E block is necessary. Therefore, a new variable $\xi^{\lambda}_{\ell_i,\ell_o}$ representing streamlined flows is introduced. The variable $\xi^{\lambda}_{\ell_i,\ell_o}$ is one when there is at least one flow entering node through port ℓ_i and leaving through port ℓ_o at wavelength λ.

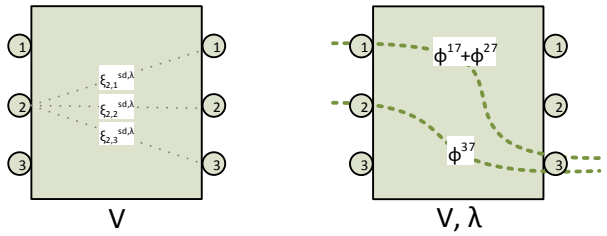


Figure 5.13: Visualisation of the MX configuration captured by $\xi^{\lambda}_{\ell_i,\ell_o}$. The flows Φ^{sd} are considered as merged compared to Figure 5.12.

The conversion of flow variables $\chi^{sd,\lambda}_{\ell_i,\ell_o}$ into stream variables $\xi^{\lambda}_{\ell_i,\ell_o}$ is realised by the big-M [32] notation bellow:

$$\sum_{\substack{sd \in V \\ sd \in \mathcal{V} \setminus \{v\}^2}} \chi^{sd,\lambda}_{\ell_i,\ell_o} \leq \xi^{\lambda}_{\ell_i,\ell_o} M \qquad v \in \mathcal{V}, \ell_o \in \omega^+(v), \ell_i \in \omega^-(v) \tag{5.26}$$

$$\sum_{\ell_i \in \omega^-(v)} \xi^{\lambda}_{\ell_i,\ell_o} \leq MF^{\lambda}_{\ell_o} + 1 \qquad v \in \mathcal{V}, \ell_o \in \omega^+(v) \tag{5.27}$$

The constraints (5.27) allow to count the number of merging flows, and the results is stored in the variable MF^{λ}_{ℓ}.

The number of wavelengths calculation

The next step on the way to find all parameters is to find the number of wavelengths participating in the contention. This is very similar approach to the MF enumeration (5.27). In the first step the indication of contention at wavelength λ must be obtained. In order to do that the big-M approach is used again:

$$MF_{\ell_o}^{\lambda} = w_{\ell_o}^{\lambda} M \qquad v \in V, \ell_o \in \omega^+(v) \tag{5.28}$$

$$\sum_{\lambda \in \Lambda} w_{\ell_o}^{\lambda} = W_{\ell_o} \qquad v \in V, \ell_o \in \omega^+(v) \tag{5.29}$$

In the second step the number of wavelength is enumerated for the output port ℓ_o.

The supplied bandwidth enumeration

All nodes in network are stable due to the constraints (5.19), so the total output port load can be obtained by simple sum of supported bandwidth routed through each output port ℓ_o:

$$\sum_{s,d \in V} \phi_{\ell_o}^{sd} = \alpha_{\ell_o} \qquad v \in V, \ell_o \in \omega^+(v) \tag{5.30}$$

O/E blocks calculation

The very last step is to calculate the number of necessary O/E blocks in order to deliver loss-less mode of operation using the (5.15). At this point, all informations were gathered together – (5.27), (5.29), (5.30). The (5.15) is then written as:

$$W_{\ell_o}.b_{0,W_{\ell_o}} + b_{1,W_{\ell_o}}.\alpha_{\ell_o} \leq OE_{\ell_o} \qquad v \in V, \ell_o \in \omega^+(v), \text{if } \alpha_{\ell_o} > b_{thresh,W_{\ell_o}}, \forall b_0, b_1 \tag{5.31}$$

$$\sum_{\ell_o \in \omega^+(v)} OE_{\ell_o} \leq OE_v \qquad v \in V \tag{5.32}$$

This formulation uses very same approach where there are dedicated variables bounded to one output port and values of these variables are summed together for each node v.

○○ The OE_v is used in the objective function (5.16) of the ILP mathematical program.

5.5.4 Parameters and variables:

In order to keep the mathematical model consistent the variables are defined here:

$$\phi_{\ell}^{sd} \geq 0, e_{sd} \geq 0 \qquad \{s,d\} \in V, \ell \in L \tag{5.33}$$

$$y_{\ell,\lambda}^{sd}, \xi_{\ell_o}^{\lambda}, w_{\ell_o}^{\lambda} \in \{0,1\} \qquad \{s,d\} \in V, \lambda \in \Lambda, \ell \in L. \tag{5.34}$$

$$MF_{\ell_o}^{\lambda}, W_{\ell_o}, OE_{\ell_o}, OE_v \in \mathbb{N}^0 \qquad v \in V, \ell_o \in \omega^+(v) \tag{5.35}$$

$$\alpha_{\ell_o} \geq 0 \qquad v \in V, \ell_o \in \omega^+(v) \tag{5.36}$$

As one can see, this ILP program relies on some variables, but considerable part of them has many parameters which makes the solution very difficult. More parameters a variable has increases the size of the variable that is represented by a matrix for the solution using the Simplex algorithm.

5.5.5 Notes to the solution

This approach suffers from scalability issue due to the complexity of used variables Section 5.5.4 and due to non-linear nature of some constraints that are linearised eventually. Another very important issue of this formulation is the dependence of parameters b_{xxx} on values of variables. This is possible to mitigate by reformulation, but this reformulation is non-linear, so this reformulation is not possible to use with linear programming approach. Additionally, some iterative heuristic can be used. In every iteration the improvement would be checked with change of routing, number of wavelengths participating in contention, and the output port ℓ_o load.

5.6 Conclusion

In this chapter the traffic properties were tackled, and the CAPEX and OPEX was formulated based on the traffic properties. The CAPEX and OPEX can be obtained quite easily from the traffic properties, but they are very problematic for the optimization analysis because they rely on the non-linear parameters which are selected based on the value of variables used for optimization. This is very crucial for mathematical program, so the conclusion is the direct CAPEX and OPEX analysis can not be carried out using the O/E analysis and mathematical program.

This approach can lead to the tangible conclusion, but using an heuristic. Using heuristic can be vital because the result can be obtained in relatively short time; however, the quality of the solution can not be guaranteed. Therefore, the viability of CAROBS must be carried in a different way.

Chapter 6

Efficient routing and wavelength assignment in CAROBS networks

The previous chapter contains the analysis of CAROBS framework behaviour and shows the direction of CAPEX and OPEXoptimizations. In order to define the optimization framework, many variables and depended constants were defined. This, however, makes the framework hard to be used for optimizations using LP. On the other hand, this framework can be used for other heuristics as description of the heuristic behaviour. Nevertheless, the focus of this chapter is on the maximal use of resources in network, and this can be only achieved using LP approach.

This section is focused on design of the accurate GRWA algorithm for CAROBS in order to show the CAROBS framework viability. To do so, the special case of the study from Section 5 is used. The special case happens when two flows are merging at node v, and one of those two flows originates at node v. This basic situation is called SLE and was recalled in Section 4.1.5. When there are flows merging as mentioned, the Phung *et al.* [44] shows that no contention occurs for classical OBS networks. The CAROBS indicates the same inclination, see Section 5, hence the focus of this section is the SLE, and based on the node properties the LP formulation is made. The LP formulation will allow only SLE merging. Barradas *et al.* [70] came up with an analysis of BLP for routing which maximizes usages of already merged flows [70], i.e., SLE maximization. His conclusion is: when WC is used the routing maximizing use of SLE performs better then the routing only based on shortest-paths. As long as CAROBS allows WC, thanks to electrical buffering, the routing maximizing SLE is worth.

When only SLE routed flow-paths are allowed, no contention will occur due to flow merging. Buffering, however, will be caused by secondary contention, i.e., when an incoming burst contents with *(i)* an aggregated burst whose header was sent towards the next node already or *(ii)* with a previously buffered burst which was buffered because

of – secondary contention. The detail description of secondary contention is addressed in Section 6.1. Interestingly, the routing maximizing SLE changes the flow-paths in network to very likely situation that has been addressed by concept called light-trails [71]. Therein, the authors studied in detail the grade of service of light-trails, and ultimately suggested integer linear programming formulation for optimizing the light-trail distribution on wavelengths. The result is a light-trail utilization of about 80%. However, there is no information about overall network efficiency, the network topology is vaguely specified, traffic requests are not specified, and it is not specified whether this 80% light-trail utilization is the maximal, mean, or median value. Therefore, the light-trail concept is very interesting, but no deeper analysis has been written, so it is very hard to compare.

The SLE maximizing routing seems to be very promising, but as been shown in [72], some buffering always prevails because of secondary contention [72], Section 6.1. Therefore, the impact of secondary contention on traffic is studied in the following text in the first place.

6.1 Aggregation Process and Contention

The traffic aggregation is a very comprehensive part of the CAROBS framework. This is very well described by Coutelen *et al.* [49]; however, a few very important details must be recalled at this point. The CAROBS framework uses AQs in order to aggregate user data for various destinations at each node. Every AQ is used to create a car for a given destination. This car is part of APs that map a few cars to be signalled by one CHP. Cars in an AP are sorted according to their distance to the destination, and all cars must be located along the same flow-path relevant to AP – closest first. Each AQ can be the head of only one AP, hence, when an AQ is full only one AP is initiated for a shipment. Some AQs, which are part of this AP, create cars and one burst train. The burst train is depicted in Figure 6.1.

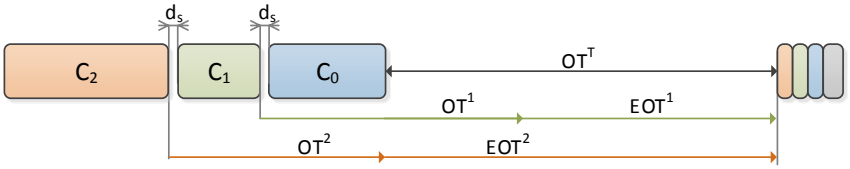


Figure 6.1: The burst train composition. Cars that transport user data are on the left side, Header is on the right side. The CHP contains the general section (gray) and the car section (colors are relevant to the cars). Each car header contains data relevant to the particular car C_x.

Every time a burst train reaches a node for which the first car C_0 is destined, the first car is disaggregated. It means the time window of the first car C_0 is available for

scheduling. This time window can be used for *a)* traffic aggregation or *b)* un-buffering of previously buffered cars, only if the new scheduling can fit into this time window, and if there is enough time to sent the CHP to reserve a flow-path. The aggregation process is depicted in Figure 6.2. This simplified figure, however, captures the SLE routing precisely because due to SLE routing flow merging never happens hence the ingress port can be in a state as depicted in Figure 6.2. There, one can see 2 burst trains preceded by their CHPs at times 1 and 5. These two CHPs are processed by the SOA Manager at CAROBS node, and reserve time windows aligned by the dash lined rectangles. These CHPs are forwarded towards to the next node at times 2 and 6. At time 3, an AQ triggers an AP initiation, and first-fit algorithm finds scheduling for the aggregating burst train. According to the distance, offset time is calculated [73] and both the CHP and the burst train are scheduled to be sent. The aggregating burst train's CHP is sent at time 4.

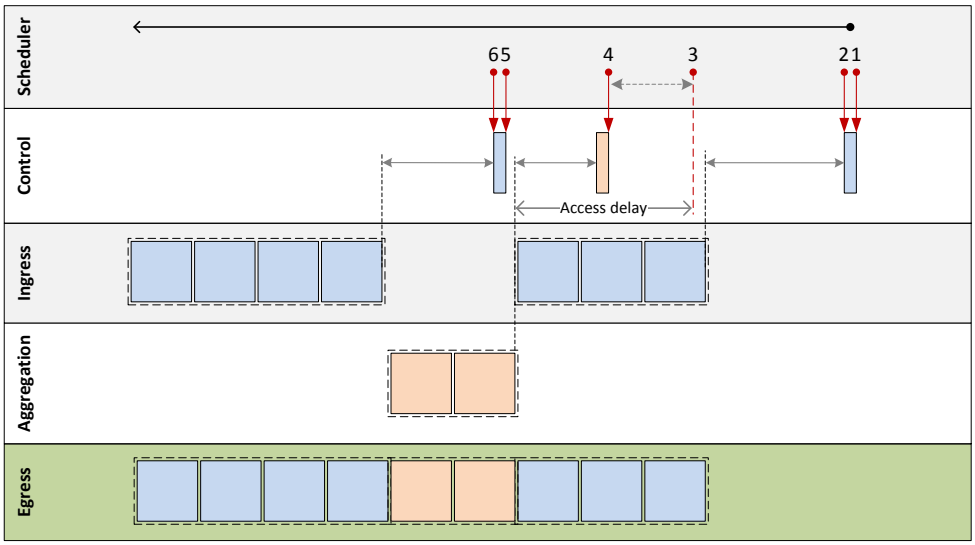


Figure 6.2: Merging of flows without any secondary contention. Blue and orange blocks represents cars. For this example, cars are of the same length, but in practice, this is not necessary.

This aggregation, however, still can cause secondary contention when the aggregation is already planned, and the new overlapping burst train's CHP arrives, as depicted in Figure 6.3. Such a situation can occur in case the incoming burst train is very close to its destination hence the offset time is small, which means the burst train arrives very soon after CHP. To be noted, this problem can cause burst train un-buffering as well. One possible resolution of this problem is variable burst train length based on empty spaces. This approach, however, does not work for buffered burst trains because they are already assembled, and their size can not be changed.

In order to avoid such unnecessary buffering, the scheduling mechanism is adjusted in order to allow aggregation/un-buffering delaying. The contenting scheduling can be

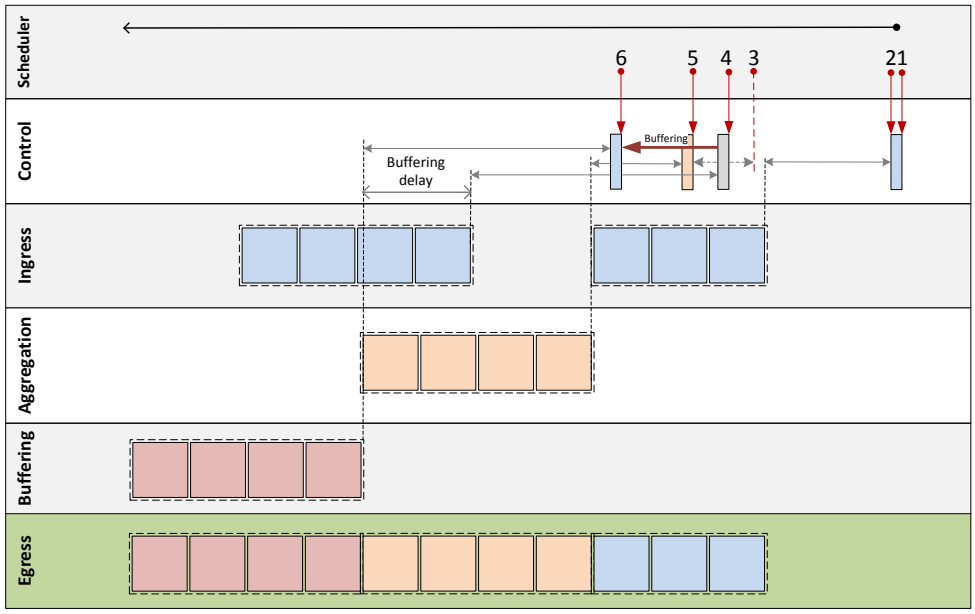


Figure 6.3: An typical situation causing secondary contention in all types of loss-less OBS paradigms.

delayed only if the CHP has not been sent. A typical situation along these lines is depicted in Figure 6.1. It is similar to the previous examples, the burst train is scheduled for sending at time 3, but at time 4, an CHP comes in for another burst train which requires an overlapping time window with the aggregating/un-buffering burst train. Such a situation can be resolved by delaying aggregating/un-buffering burst train, so the CHP is not send at time 6, as it was originally planned, but at time 7. This rescheduling algorithm workflow is described in Algorithm 1 in more details.

Although the delaying mechanism can resolve some secondary contention there is still a way how the secondary contention can occur. This situation is depicted in Figure 6.5. The secondary contention occurs when the aggregating/un-buffering burst train's CHP has been already sent by the time a CHP for an incoming burst train arrives to SOA Manager, moment 6. In this situation the burst train on ingress can only be buffered in loss-less paradigm.

6.1.1 Aggregation performance

The SLE routing performance in CAROBS networks is very hard to quantify by analytical tools. The M/M/k/k model does not take into account SLE [44], and the modification of M/M/k/k for SLE does not work for the CAROBS framework because the CAROBS framework uses buffering [72]. Buffering produces additional offered load to the node in CAROBS architecture, and this additional offered load is proportional to the node offered load on the optical domain. This additional offered load is added to the offered load from

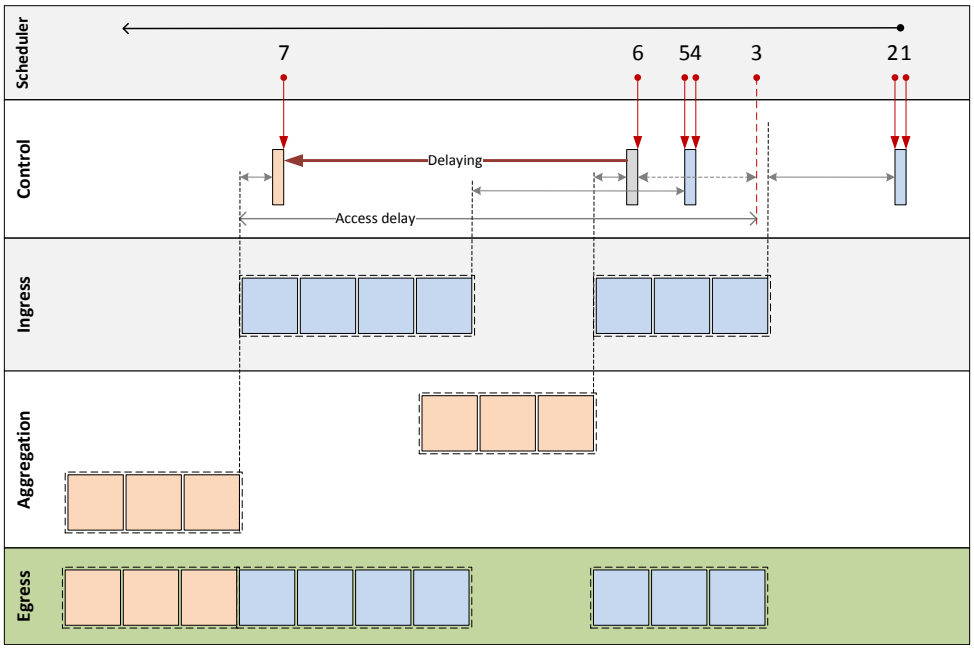


Figure 6.4: An typical situation how the proposed delaying mechanism can resolve secondary contention.

adjacent networks. Subsequently, the node offered load in the electrical domain α can be quantify as follows:

$$\alpha = \alpha_{\text{agg}} + \text{BBP}_i \times \alpha_{\text{input}} \tag{6.1}$$

where α_{agg} is the offered load from adjacent networks, α_{input} is the offered load on the optical domain, and most importantly BBP is buffering probability. All the buffered traffic is unbuffered after a while producing load $BBP_i * \alpha_{\text{input}}$. The BBP, however, has not been analytically quantified. To be noted, this offered load does not represent the total node offered load, but only the load of flows that take the same egress (combination of port and wavelength).

According to the authors' knowledge, there is no analytical formulation of BBP. Hence, the BBP is analyzed through simulation. Objective is to analyze the dependence of BBP, BD, access delay (ACC) and secondary contention ratio (SECR) on location i and level of offered load. The motivation for this analysis is to find limits of SLE routing in CAROBS framework based networks. Analysis using OMNeT++ simulator and CAROBS framework models [53] is performed. The input traffic from adjacent networks has Poisson characteristics, and the bandwidth data rate is *a)* 10 Gbps, and *b)* 100 Gbps. The simple network topology depicted in Figure 6.6 is used. Therein, three location scenarios are captured, and these locations are iteratively verified. In the experiments, two flows are used. Flow $0 \rightarrow 4$ is used in all scenarios, and this flow represents all the flows which have been merged into one flow in one of the previous nodes before node C_0.

Such a part of the network, however, is not depicted in Figure 6.6.

Algorithm 1 Rescheduling algorithm

```
1: CHP is processed by a SOA Manager
2: p ← Ingress port of incoming burst train
3: λ ← Wavelength of incoming burst train
4: t_ARR ← Burst train arrival time
5: Δ ← Burst train duration
6: if Contention(p, λ, t_ARR, Δ) then
7:     CanReschedule ← True
8:     for OverlappingSchedulings_p^λ → s do
9:         if Header of s was sent then
10:            CanReschedule ← False
11:    if CanReschedule then
12:        Reserve λ from t_ARR for Δ
13:        for Scheduling_p^λ → s do
14:            RecalculateWaiting(s)
15:    else
16:        Burst train is buffered.
17: else
18:    Reserve λ from t_ARR for Δ
```

The node load in this performance analysis is given by the combination of offered load of flow $0 \rightarrow 4$ which offers load b_1, and the second flow A), B), and C) which offers load b_2 then the node load is defined as (6.2).

$$\textsc{Load} = \frac{b_1 + b_2}{B}, \tag{6.2}$$

where B represents the wavelength bandwidth, $b_i, i \in \{1, 2\}$ represent the data rate of contributing flows such that $\textsc{Load} \in [0, 1]$. The inter-arrival time between cars is random but proportional to the load of participating flows $b_i, i \in \{1, 2\}$, thus the simulation are repeated multiple-times in order to achieve results that are not correlated to one combination of parameters $b_i, i \in \{1, 2\}$. Subsequently, simulations are run for every selected level of Load 50 times with different traffic pattern. The dependence of performance parameters on the Load, and confidence intervals of obtained data are depicted in Figure 6.7.

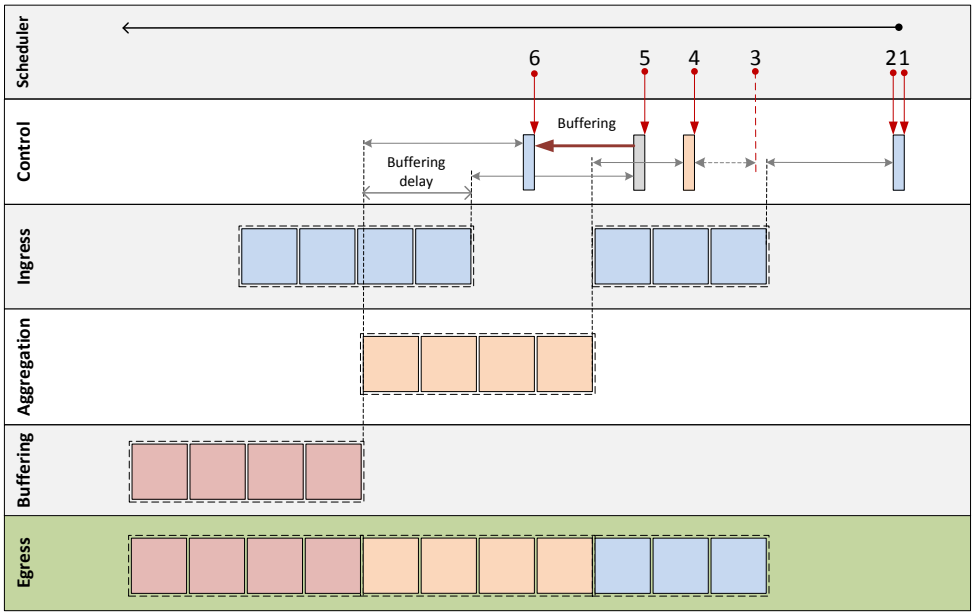


Figure 6.5: An typical situation when secondary contention can occur; although, delaying mechanism is used.

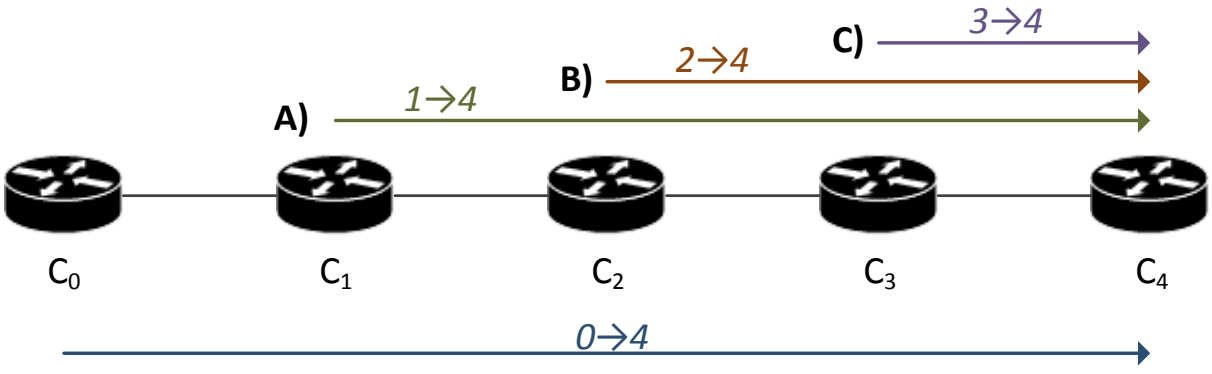


Figure 6.6: Topology used for testing the dependence of performance parameters on the location where secondary contention occurs. The test consists of three scenarios A), B), and C) such that contentions only can happen on nodes C_1, C_2, and C_3 respectively.

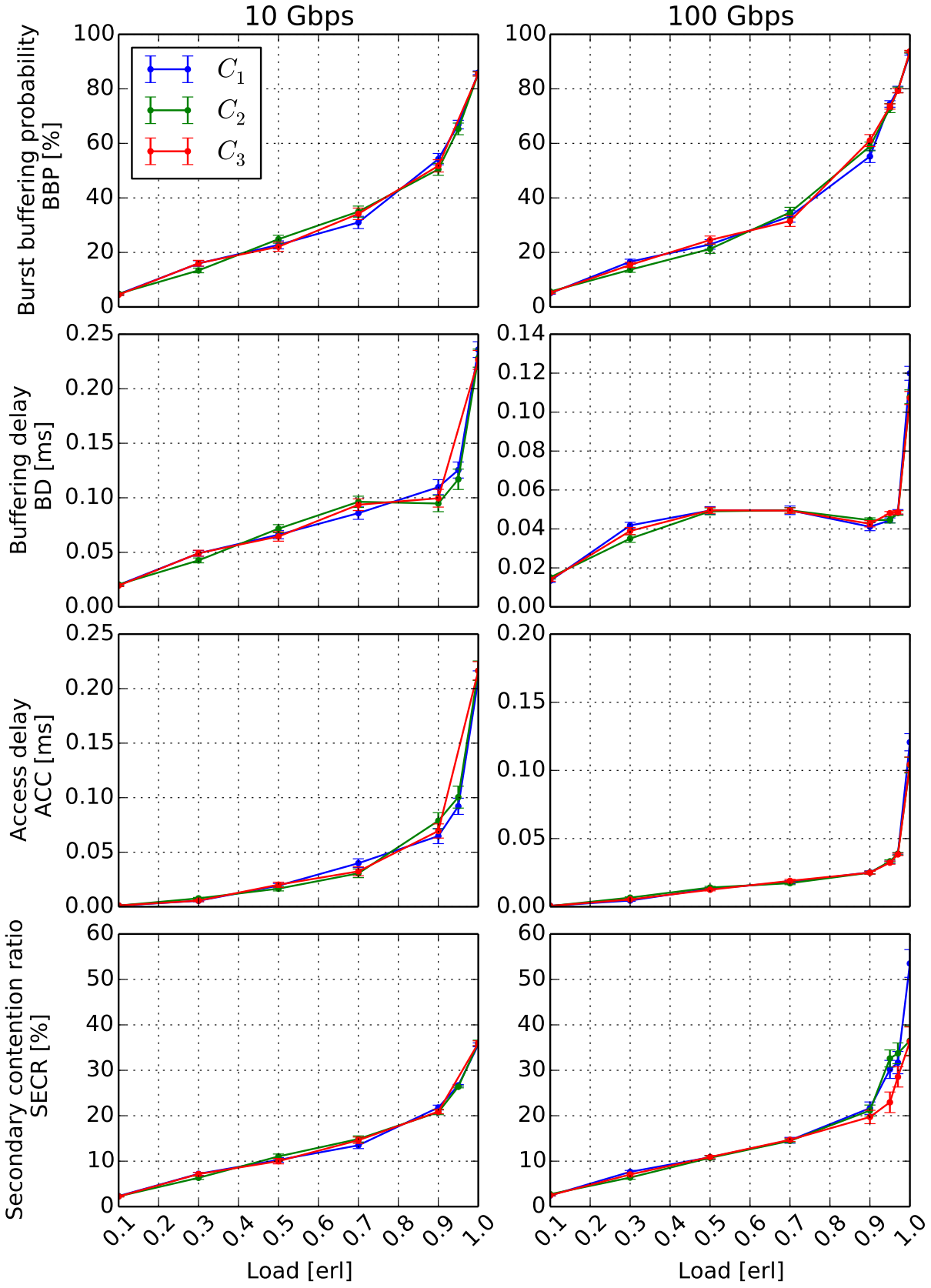


Figure 6.7: Comparison of results for 10Gbps and 100Gbps wavelengths. Each line is identified by the node name where the performance parameter was measured.

One can see that all considered performance parameters do not depend on a specific location in the network, see Figure 6.7. These results are direct contradiction to assumptions used as the principal pillars of light-trails [71], [74]. The location independence is a very important property because it allows modelling of SLE routing in CAROBS networks without implementation of very special constraints. Additionally, one can see the exponential increase in all monitored parameters when the Load 0.9 erl is crossed. This is the buffering effect. More importantly, one can see that BBP does not reach the 100% value for 1 erl, which means the node remains stable. Node stability was also verified using stationary test based on linear regression test of BD [72].

In other words, if the BD does not increase in time then it can be claimed that there is a certain maximum size of electrical buffer that will never be exceeded, and that the waiting in electronic buffer has a limit. This very interesting observation provides the SECR that gives information of how much buffering was caused by previously un-buffered traffic, and the $(1 - \text{SECR})$ gives sense how much of buffering was caused by aggregation. Results show relatively high end to end delays (E2Es) an overloaded network with more hops. Nevertheless, the E2E is less than 1 ms even for networks with 10 hops at wavelength bandwidth 10 Gbps, and this end-to-end delay is not possible to achieve with current technology.

One possible way of improving the performance parameters is by adding extra capacity to channels. In this study, no extra capacity to the channel was added. On the other hand, if this exponential exponential increase of performance parameters values: BD, ACC, BBP, and SECR was meant to be eliminated then the wavelength capacity could be simply increased. This claim is proven by small experiment bellow where the transport capacities of all the links in the network was increased about 3%. Results are depicted in Figure 6.8. One can see that the exponential increase of performance parameters was effectively suppressed. The behaviour for the three location scenarios is not depicted because the difference was negligible in Figure 6.8, so parameters for from all locations were put together, mean value was calculated, and this mean value is depicted in Figure 6.8. Interesting is the independence of probability parameters on wavelength data rate. Whereas, the time domain performance parameters: BD, ACC get significantly improved. This improvement is as expected, for the burst trains contain more packets within triggering intervals, thus the burst train length/duration tends to be deterministic as AQ fill faster.

It is worth noting that BD is considerably low. Considering that up to 90% of burst trains where each train transmits hundreds of user's packets, is delayed by $50\mu s$ when the bypassing node is fully loaded at datarate 100 Gbps, then one can see that the CAROBS framework performance is worth of further research.

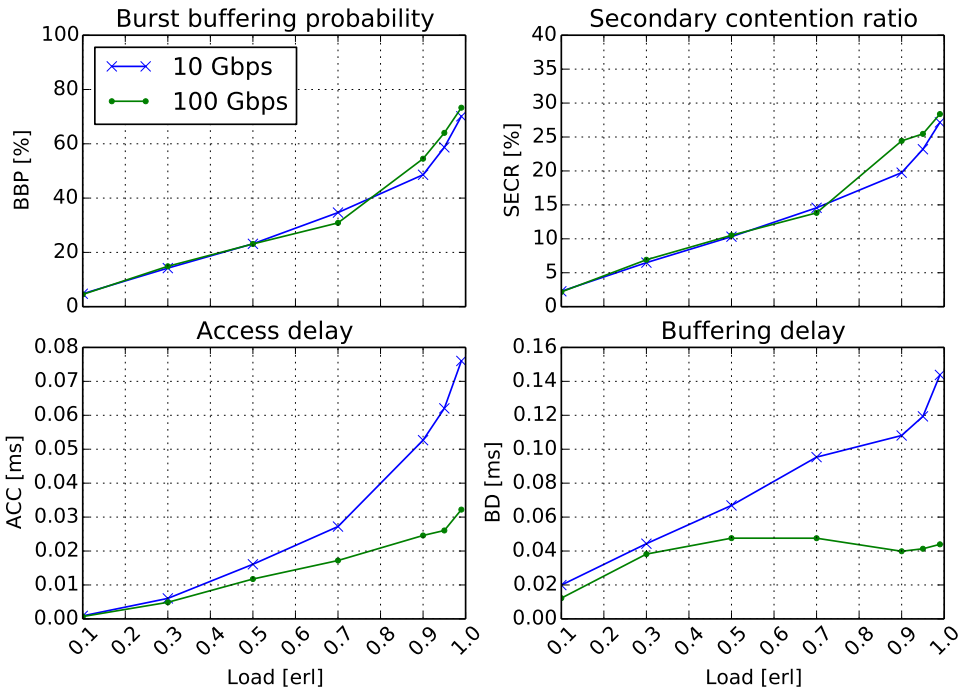


Figure 6.8: Channel maximal capacity increased about 0.3erl. Only one line for each datarate is provided because of independence on the location where the performance parameters were measured. These values correspond to the average values from Figure 6.7.

In future work, it will be very important to bear in mind the maximal efficiency of transmission.

6.2 Stream-line effect based traffic routing

The previous section evaluated various network performance parameters on a small trial network, under the assumption of a SLE routing within the CAROBS framework. Results show a lower bound of performance parameters in the context of further studies and representative network topologies. In order to design a CAROBS network supporting the maximum possible traffic for a given network topology and minimal impact on the performance parameters, a new routing approach called SLE-RWA using only SLE for traffic aggregation is defined in this section. To keep minimal values of monitored performance parameters and achieve high network efficiency techniques of operational research (OR) such as LP were used.

In the sequel of this section, first the SLE-RWA routing and provisioning problem with a compact ILP model is defined. In order to overcome the scalability issues of that first ILP formulation, a decomposition ILP model relying on a column generation formulation is investigated.

6.2.1 Compact SLE-RWA ILP Formulation

The ILP formulation of SLE-RWA is designed to maximize the traffic D_{sd} deployment in a given network $G(\mathcal{V}, \mathcal{L})$. The objective function of the ILP SLE-RWA is to minimize amount of unsupported traffic U_{sd} between each node pair $(v_s, v_d) \in \mathcal{SD}$:

$$\min \sum_{(v_s,v_d)\in\mathcal{SD}} U_{sd}. \tag{6.3}$$

This objective function is subject to two set of constraints that a) tackle the traffic routing b) ensure no merging flows in the network. In this section constraints are only categorized because the detail description is provided in the Section 6.2.3, and the goal of this section is to only give the sense of the problem and emphasis on the motivation why the SLE-RWA using a decomposition technique is formulated.

The first set of constraints of ILP SLE-RWA focuses on the traffic routing in the network. The ILP SLE-RWA allows to groom more traffic flows onto the same wavelength if the wavelength capacity B is not exceeded. The un-supported traffic U_{sd} is calculated using constraints (6.5).

$$\sum_{\ell\in\omega^+(v)} \varphi^{sd}_{\ell,\lambda} - \sum_{\ell\in\omega^-(v)} \varphi^{sd}_{\ell,\lambda} = \begin{cases} b^{sd}_{\lambda} & \text{if } v = v_s \\ -b^{sd}_{\lambda} & \text{if } v = v_d \\ 0 & \text{otherwise} \end{cases}$$

$$v \in V, (v_s, v_d) \in \mathcal{SD} \tag{6.4}$$

$$D_{sd} = \sum_{\lambda\in\Lambda} b^{sd}_{\lambda} + U_{sd} \qquad (v_s, v_d) \in \mathcal{SD} \tag{6.5}$$

$$\sum_{(v_s,v_d)\in\mathcal{SD}} \varphi^{sd}_{\ell,\lambda} \leq B \qquad \ell \in L, \lambda \in \lambda. \tag{6.6}$$

Every request D_{sd} is routed by one flow-path at each wavelength. Additionally, the ILP SLE-RWA formulation tackles loops at intermediate nodes, so efficient usage of links in

the network can be achieved: For $\lambda \in \lambda, (v_s, v_d) \in \mathcal{SD}$,

$$b_{\min}\, y^{sd}_{\ell,\lambda} \leq \varphi^{sd}_{\ell,\lambda} \leq D_{sd}\, y^{sd}_{\ell,\lambda} \qquad \ell \in L \tag{6.7}$$

$$\sum_{\ell \in \omega^+(v)} y^{sd}_{\ell,\lambda} \leq 1; \sum_{\ell \in \omega^-(v)} y^{sd}_{\ell,\lambda} \leq 1 \qquad v \notin \{v_s, v_d\} \tag{6.8}$$

$$\sum_{\ell \in \omega^-(v_s)} y^{sd}_{\ell,\lambda} = \sum_{\ell \in \omega^+(v_d)} y^{sd}_{\ell,\lambda} = 0 \tag{6.9}$$

$$\sum_{\ell \in \omega^+(v_s)} y^{sd}_{\ell,\lambda} = \sum_{\ell \in \omega^-(v_d)} y^{sd}_{\ell,\lambda} \leq 1. \tag{6.10}$$

The SLE is ensured in the network by following set of constraints:

$$y_{\ell_i,\ell_o,\lambda} \geq y^{sd}_{\ell_i,\lambda} + y^{sd}_{\ell_o,\lambda} - 1 \qquad \lambda \in \lambda, (v_s, v_d) \in \mathcal{SD}, v \in V \setminus \{v_s, v_d\}$$
$$\ell_i, \ell_o \in L : \text{DST}(\ell_i) = \text{SRC}(\ell_o) = \{v\} \tag{6.11}$$

$$\sum_{\ell_i \in L : \text{DST}(\ell_i) = \text{SRC}(\ell_o)} y_{\ell_i,\ell_o,\lambda} \leq 1 \;\; \ell_o \in L, \lambda \in \Lambda, \tag{6.12}$$

which express the wavelength connectivity at each node $v \in \mathcal{V}$. Every wavelength connectivity is captured by the variable $y_{\ell_i,\ell_o,\lambda}$ which is one only if there is at least one $(v_s, v_d) \in \mathcal{SD}$ flow going from input port connected to link ℓ_i to output port connected to link ℓ_o, routed on wavelength λ. The difficulty to solve the routing for the ILP SLE-RWA comes from the huge number of variables that are necessary:

$$y^{sd}_{\ell,\lambda} \in \{0, 1\} \qquad \ell \in L, (v_s, v_d) \in \mathcal{SD}, \lambda \in \Lambda \tag{6.13}$$

$$y_{\ell_i,\ell_o,\lambda} \in \{0, 1\} \qquad \ell_i, \ell_o \in L : \text{DST}(\ell_i) = \text{SRC}(\ell_o), \lambda \in \Lambda \tag{6.14}$$

$$\varphi^{sd}_{\ell\lambda} \geq 0 \qquad \ell \in L, \lambda \in \Lambda, (v_s, v_d) \in \mathcal{SD} \tag{6.15}$$

$$b^{sd}_{\lambda} \geq 0 \qquad \lambda \in \Lambda, (v_s, v_d) \in \mathcal{SD} \tag{6.16}$$

$$U_{sd} \geq 0 \qquad \{s, d\} \in \mathcal{SD}. \tag{6.17}$$

Most variables are indexed by wavelength λ, but their value is nonzero for only one wavelength λ. This characteristic of the ILP SLE-RWA model generates scalability issues even for small networks with a large number of wavelengths. Using a decomposition model (see Section 6.2.2) allows solving each set of constraints for one generic wavelength as an independent ILP program, consequently with significantly less variables. In other words, using a decomposition model, one large complex problem is not solved, but a number of significantly less complicated and smaller problems [75] are solved.

6.2.2 A SLE-RWA Decomposition Model

The SLE-RWA column generation model relies on the selection of a set of routing and provisioning configurations, which are dynamically generated. Each configuration is defined by a set of provisioned flow-paths, which are all using the same wavelength. Flow-paths correspond to SLE compliant paths, with each path carrying a fraction of D_{sd} traffic for a given pair of source and destination nodes $(v_s, v_d) \in \mathcal{SD}$. The concept of provisioned configurations is depicted in Figure 6.9.

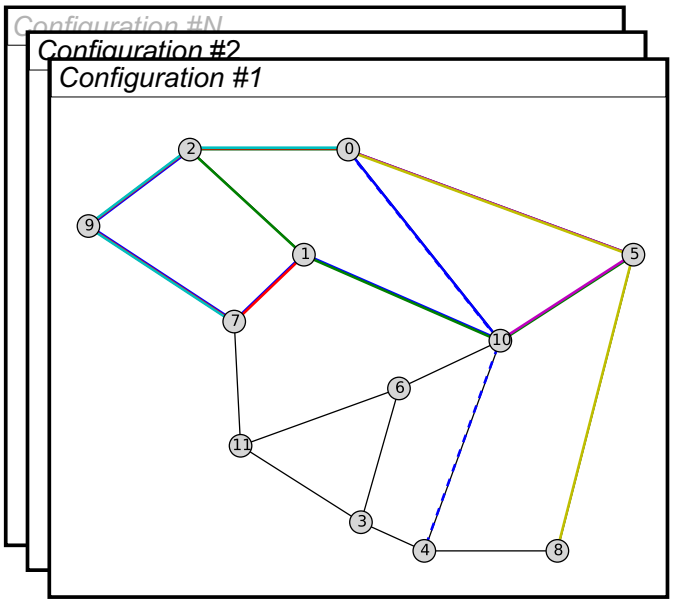


Figure 6.9: An example of a configuration set containing N configurations. Paths are depicted by various colours in this example; however, colours does not imply different wavelengths because all paths p^c are assigned the same wavelength.

This concept relies on the parameters and the variables of the column generation model.

$\mathcal{SD}$ – set of node pairs (v_s, v_d) such that $D_{sd} > 0$

C – overall set of wavelength configurations, indexed by c

c – generic wavelength transmission: it is characterized by a set of SLE compliant paths p^c, i.e., there are no paths merging at one of their intermediate node. Each path is characterized by

- p^c_ℓ =1 if link ℓ belongs to the path, 0, otherwise,
- b^c_{sd} – amount of bandwidth routed in c for $(v_s, v_d) \in \mathcal{SD}$,
- $a^c_{sd} = 1$ if there exists at least one $v_s \rightsquigarrow v_d$ flow in configuration c,

z_c – a non-negative integer variable specifying the number of wavelengths for which configuration c is selected

The objective function is written as follows. In order to break the ties efficiently, the initial objective function (6.3) was refined:

$$\min \sum_{(v_s,v_d)\in\mathcal{SD}} U_{sd} + \text{PENAL}_\text{W} \times \sum_{c\in C} z_c + \text{PENAL}_\text{P} \times \sum_{c\in C} \left(\sum_{(v_s,v_d)\in\mathcal{SD}} a_{sd}^c \right) z_c. \tag{6.18}$$

The primary objective is to grant all the requests and therefore to minimize the amount of bandwidth that is not provisioned. The second objective minimizes the number of wavelengths (i.e., $\sum_{c\in C} z_c$) and maximizes the bandwidth usage (i.e., $\sum_{c\in C} \left(\sum_{(v_s,v_d)\in\mathcal{SD}} a_{sd}^c \right) z_c$). Note that, throughout the minimization of $\sum_{c\in C} \left(\sum_{(v_s,v_d)\in\mathcal{SD}} a_{sd}^c \right) z_c$, the number of routes is minimized as well, and consequently maximizes the utilization of the established routes. Penalties may become necessary for some terms of the objective if the value of $\sum_{(v_s,v_d)\in\mathcal{SD}} U_{sd} \neq 0$ in order to make sure of the sub-objective priorities. In the experiment bellow the values or penalties are $\text{PENAL}_\text{W} = \text{PENAL}_\text{P} = 1$ because presumably there is always enough wavelengths in order to grant all the requested traffic.

Selection of the best routing and provisioning configurations is done thanks to the following set of constraints.

$$\sum_{c\in C} z_c \leq W \tag{6.19}$$

$$\sum_{c\in C} d_{sd}^c z_c \geq D_{sd} - U_{sd} \qquad (v_s, v_d) \in \mathcal{SD} \tag{6.20}$$

$$z_c \in \text{Z}^+ \qquad c \in C. \tag{6.21}$$

Constraints (6.19) ensure that number of available wavelengths is not exceeded. Constraints (6.20) correspond to the demand constraints: subject to (6.19), the aim is to maximize the grade of service (GoS). In the sequel, refers (6.18) - (6.21) to the restricted master problem (RMP) of the column generation (CG) SLE-RWA model.

6.2.3 Solution of the CG SLE-RWA Model

In order to solve the CG SLE-RWA model, first column generation method is used in order to solve its continuous relaxation (i.e., with $z_c \geq 0$ instead of $z_c \in \text{Z}^+$, and then a branch-and-bound algorithm (ILP solver of Cplex [37]) in order to derive an integer solution. In order to solve the linear relaxation of the CG SLE-RWA model a so-called RMP is derived from the master problem (MP), which contains a restricted subset of configurations. The column generation algorithm (see, e.g., [32] if not familiar with column generation) consists

in solving alternatively the RMP and the so-called Pricing Problem, which dynamically generates new augmenting configurations, until the optimality condition is satisfied. An augmenting configuration is such that, if added to the current RMP, allows improving the value of its objective. It is characterized by a negative reduced cost, i.e., the objective of the pricing problem (see again [32] for its definition). Optimality condition is satisfied when no new configuration with a negative reduced cost can be built. The flowchart capturing the iterative process of the column generation is depicted in the upper part of Figure 6.10.

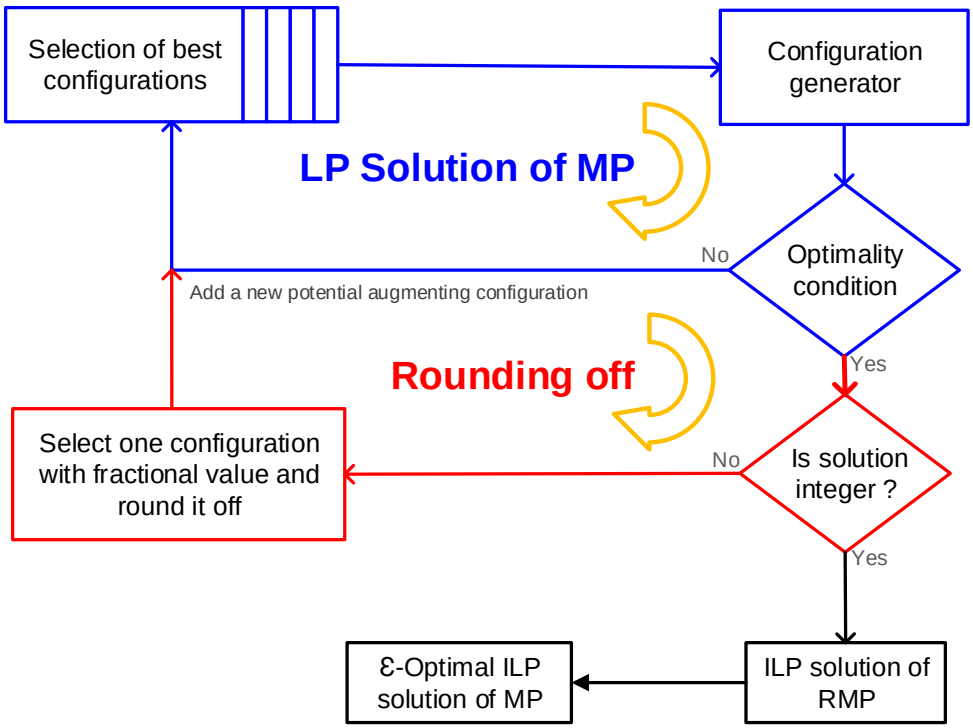


Figure 6.10: Flowchart of the Solution Process.

Once the optimality condition is satisfied, the optimal value $z^{\star}_{\text{LP}}$ of the linear relaxation of (6.19) - (6.21) has been reached. The focus shifts to the derivation of an ILP solution. In order to do so, *rounding off* approach is used in this experiment. Routing off consists of iteratively rounding steps of some selected fractional z_c variables to an integer value and then re-optimizing the LP solution of the resulting RMP. The process is iterated until all z_c variables have an integer value. Subsequently, in practice, it tentatively allows reaching an integer solution that can be very close to the optimal integer value, or at least an integer value $\tilde{z}_{\text{ILP}}$ for which accuracy can be estimated (denoted by $\varepsilon = (\tilde{z}_{\text{ILP}} - z^{\star}_{\text{LP}})/z^{\star}_{\text{LP}}$).

6.2.4 Dynamic Generation of an Augmenting Configuration

The *configuration generation* is a key step in the SLE-RWA solution process as it allows the generation of SLE compliant improving configurations through communication with

the RMP [32], [66]. Each path in the network is calculated using flow constraints, and SLE enforcing constraints. Variables of the configuration generation problem, i.e., the so-called pricing problem (PP) in the Operations Research literature, are as follows. (c index is omitted to alleviate the presentation of the PP).
b_{sd} – amount of routed bandwidth in the configuration from node v_s to v_d,
$\varphi_\ell^{sd} \in \mathbb{R}^+$ – amount of routed bandwidth in the configuration from v_s to v_d assigned to link ℓ,
y_ℓ^{sd} – decision variable which is $y_{sd}^\ell = 1$ if some traffic from v_s to v_d is routed on link ℓ in the configuration,
$y_{\ell_i \ell_o}$ – decision variable indicating that $(v_s, v_d) \in \mathcal{SD}$ flow is routed from ℓ_i to ℓ_o at node $v = \text{SRC}(\ell_o)$.

The objective function of the pricing problem, i.e., the reduced cost is defined as:

$$\overline{\text{COST}} = 1 + \sum_{(v_s,v_d)\in\mathcal{SD}} a_{sd} - u^{(6.19)} - \sum_{(v_s,v_d)\in\mathcal{SD}} u_{sd}^{(6.20)}\, b_{sd} \tag{6.22}$$

where $u_{sd}^{(6.20)}$ is the value of the dual variables associated with constraints (6.20) and $u^{(6.19)}$ is the value of the dual variables associated with constraints (6.19). The values of the dual variables are the corner stone of the alternate solution of the RMP and PP in the column generation algorithm.

Constraints are written as follows. They express the constraints in order to build a routing and provisioning configuration.

$$\sum_{\ell\in\omega^+(v)} \varphi_\ell^{sd} - \sum_{\ell\in\omega^-(v)} \varphi_\ell^{sd} = \begin{cases} b_{sd} & \text{if } v = v_s \\ -b_{sd} & \text{if } v = v_d \\ 0 & \text{otherwise} \end{cases}$$

$$v \in V, (v_s, v_d) \in \mathcal{SD} \tag{6.23}$$

$$y_\ell^{sd} + y_{\text{OPP}(\ell)}^{sd} \leq 1 \qquad \ell \in L, (v_s, v_d) \in \mathcal{SD} \tag{6.24}$$

$$\sum_{(v_s,v_d)\in\mathcal{SD}} \varphi_\ell^{sd} \leq B \qquad \ell \in L \tag{6.25}$$

$$b_{\min}\, y_\ell^{sd} \;\leq\; \varphi_\ell^{sd} \leq D_{sd}\, y_\ell^{sd} \qquad \ell \in L, (v_s,v_d) \in \mathcal{SD} \tag{6.26}$$

$$b_{\min}\, a_{sd} \;\leq\; b_{sd} \leq D_{sd}\, a_{sd} \qquad \ell \in L, (v_s,v_d) \in \mathcal{SD} \tag{6.27}$$

$$\sum_{\ell\in\omega^-(v_s)} y_\ell^{sd} = \sum_{\ell\in\omega^+(v_d)} y_\ell^{sd} = 0 \qquad (v_s,v_d) \in \mathcal{SD} \tag{6.28}$$

$$\sum_{\ell\in\omega^-(v)} y_\ell^{sd} = \sum_{\ell\in\omega^+(v)} y_\ell^{sd} \leq 1 \qquad v \in V : v \notin \{v_s, v_d\}$$

$$(v_s,v_d) \in \mathcal{SD} \tag{6.29}$$

$$\sum_{\ell\in\omega^-(v_d)} y_\ell^{sd} \leq 1; \sum_{\ell\in\omega^+(v_s)} y_\ell^{sd} \leq 1; \quad (v_s,v_d) \in \mathcal{SD} \tag{6.30}$$

$$y_{\ell_i}^{sd} + y_{\ell_o}^{sd} - 1 \leq y_{\ell_i\ell_o} \qquad \ell_i, \ell_o \in L : \textsc{dst}(\ell_i) = \textsc{src}(\ell_o) = \{v\},$$

$$v \in V, (v_s,v_d) \in \mathcal{SD} : v \notin \{v_s, v_d\} \tag{6.31}$$

$$\sum_{\ell_i\in L:\textsc{dst}(\ell_i)=\textsc{src}(\ell_o)} y_{\ell_i\ell_o} \leq 1 \qquad \ell_o \in L \tag{6.32}$$

$$y_\ell^{sd} \in \{0,1\} \qquad \ell \in L, (v_s,v_d) \in \mathcal{SD} \tag{6.33}$$

$$d_{sd} \geq 0 \qquad (v_s,v_d) \in \mathcal{SD} \tag{6.34}$$

$$\varphi_\ell^{sd} \geq 0 \qquad (v_s,v_d) \in \mathcal{SD}, \ell \in L \tag{6.35}$$

$$y_{\ell_i\ell_o} \in \{0,1\} \qquad \ell_i, \ell_o \in L : \textsc{dst}(\ell_i) = \textsc{src}(\ell_o). \tag{6.36}$$

Constraints (6.23), also called flow constraints, allow computing the amount of bandwidth that can be routed from v_s to v_d in the configuration under construction; subsequently, these constraints allow finding a path through the given network. The resulting flow-path is not necessarily the shortest, and does not use full capacity of each link $\ell \in L$ because grooming is allowed. This means more flows can use the same link ℓ. In other words, one traffic request D_{sd} can be routed on multiple wavelengths, i.e., different links.

Constraints (6.24) prevent isolated loops ($\textsc{opp}(\ell)$ is the link in the opposite direction of ℓ: if $\ell = (v, v')$, then $\textsc{opp}(\ell) = (v', v)$). For instance, without such constraints, nothing prevents an isolated loop from v to v', and then from v' to v. Another isolated loop with three nodes, would be a loop from $v_1 \rightsquigarrow v_2 \rightsquigarrow v_3 \rightsquigarrow v_2 \rightsquigarrow v_1$. Such isolated loops do not violate any constraint and can only be eliminated thanks to constraints (6.24).

However, the capacity of each link can not be exceeded, hence the constraints (7.8). Constraints (6.26) establish the relations between the y_ℓ^{sd} and the φ_ℓ^{sd} variables, i.e., if the flow is equal to 0 on ℓ for a given source and destination pair of nodes, then y_ℓ^{sd} must be set to 0, and vice versa, if $y_\ell^{sd} = 0$, then the flow must be null on ℓ for the (v_s, v_d) node

pair, under the assumption that $b_{\min} > 0$. On the other hand, if $y_\ell^{sd} = 1$, i.e., ℓ belongs to one of the routes selected for provisioning the D_{sd} demand, then the flow value must lie between $b_{\min}$ and D_{sd} (as long as the link capacity constraints are satisfied). Similarly, constraints (6.27) establish the relations between the a_{sd} and b_{sd} variables.

The variables y_ℓ^{sd} are very important because allow to adjust the path selected by (6.23). Then SLE in traffic routing. In the first step, however, are these variables used for loop prevention. The loops are very important to reduce because they can cause configuration c underutilization, and decrease of all monitored performance parameters. The loops occur because precalculated paths are not used. Constraints (6.28) and (7.11) prevent loops at the source, destination and intermediary nodes, respectively in the pricing of the SLE-RWA.

At last but not least the PP contains constraints enforcing SLE in the configuration under construction. The SLE is enforced using two set of constraints (7.12) and (6.32) where the constraints (7.12) only generate the temporary OXC configuration, i.e., connection matrix, by generating the variables $y_{\ell_i\ell_o}$. The variables are $y_{\ell_i\ell_o} = 1$ only if there is a flow routed from link ℓ_i to link ℓ_o. These constraints take into account already streamlined flows coming from input link ℓ_i. The constraints (7.12) can be written in detail as:

$$y_{\ell_i\ell_o} = \max_{(v_s,v_d)\in\mathcal{SD}:v\notin\{v_s,v_d\}} y_{\ell_i}^{sd}\, y_{\ell_0}^{sd}$$

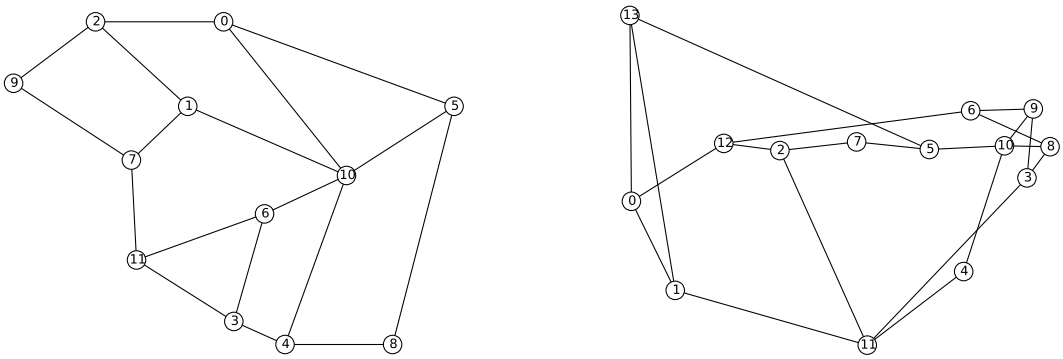
$$v \in V\ ;\ \ell_i,\ell_o \in L : \ell_i \cap \ell_o = \{v\}.$$

This formulation means that if there is at least one flow routed from ℓ_i to ℓ_o then $y_{\ell_i\ell_o} = 1$. Subsequently, the constraints (6.32) enumerate the number of flows routed to output link ℓ_o of node $v = \text{SRC}(\ell_o)$. Setting the upper limit to be one any merging is forbidden, which means only aggregation at node $v = \text{SRC}(\ell_o)$ is allowed. In other words the SLE is enforced in the network. Remaining constraints take care of the domains of the variables.

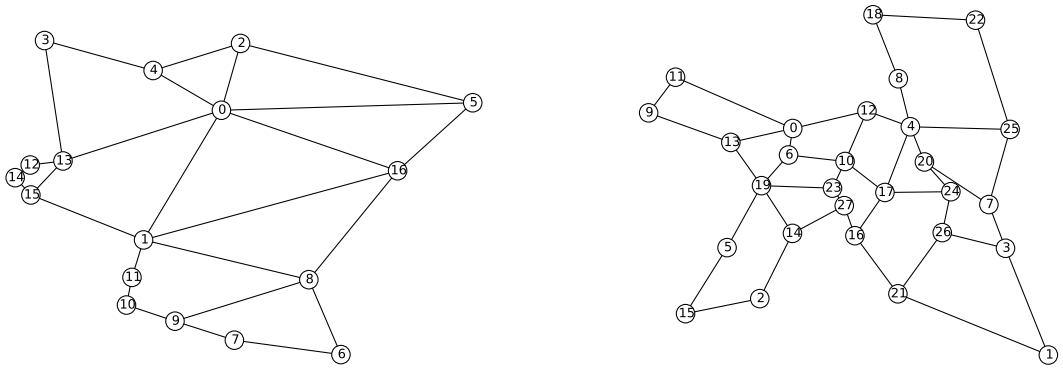
6.3 Experiments

6.3.1 Experiment Settings

The SLE-RWA optimization algorithm described in the previous section provides routing with no merging flows, but still with some contentions caused by the aggregations – secondary contention as described in Section 6.1. Although, the buffering is seen beneficial in terms of burst loss avoidance, the buffering introduces extra delay to every buffered burst train. Generally, the buffering is a new term in all-optical networks; therefore, it is



(a) Polska (b) Nobel



(c) Germany (d) Nobel-EU

Figure 6.11: Network topologies used for analysis, taken from SNDlib [76]

important to verify the impact of the traffic and network topology on CAROBS network performance in terms of SLE-RWA routing. For the purpose of this study, the CAROBS network performance is quantified by BBP, ACC, E2E, and BD. These performance parameters describe CAROBS nodes in time domain, and reflect the network behaviour from an ISP perspective. The time domain analysis is carried out using simulation models written using OMNeT++ [53].

In order to verify the CAROBS network performance, tests are run on four network topologies depicted in Figure 6.3.1. These topologies come from SNDlib [76]; however, since the traffic D_{sd} provided by SNDlib was not sufficient for the purpose of experiments, the D_{sd} was generated using random number generator based on uniform distribution.

Based on the routing obtained from SLE-RWA 20 simulations were carried out to mitigate the impact of random number generator seed number used in the simulator for traffic generation. The simulations are run using OMNeT++ simulator and CAROBS models [53]. The source nodes s supply traffic generated according to a Poisson distribution to the CAROBS network in the simulation. The source nodes generate payload packets with constant size (100 kb). Then these packets create flows to supply aggregation queues to generate bursts. JET [13] is used as a signaling protocol and LAUC-VF algorithm [68] for burst assembly. Additionally, the rescheduling mechanism discussed in Section 6.1 was

used. The data rate of each wavelength is 10 Gbps, however, as one can see in the Section 6.1.1 the increase of wavelength data rate only improves the ACC and BD. In other words this means that even E2E would improve, but the BBP would not change. The reason to use 10 Gbps is the simulation duration, for the simulations using 100 Gbps wavelength data rate last significantly longer.

6.3.2 Simulation Results

The performance parameters for the four topologies of CAROBS networks are captured in Tab. 6.1. Values of performance parameters are calculated as a mean value over all nodes in the network where these performance parameters are measured. Information about confidence intervals are not provided because the values of Standard error are less than 5% of mean value hence providing such a small number would impair the table readability.

The simulations are carried out for traffic D_{sd} that SLE-RWA distributes on multiple wavelengths and flow-paths. In other words the network load is already given, so the load does not change in this analysis. The network load is quantified as *Average link utilization* which captures the original (6.1) in context of a network. The (6.1) represents load offered to one outgoing port, i.e., outgoing link in the Section 6.1.1 hence for more links the average value of load supported by all links is calculated.

When one look at the values of performance parameters of the Polska network, one can see that the value of BBP and ACC highly correlate with the simple case discussed in Section 6.1.1 at the proportional level of LOAD. The BD is higher which is fine as long as it is known that the values from Section 6.1.1 are the lower bounds. The E2E is considerably higher which is caused by the bigger topology than what was used in the simple case. One can see that the burst train is buffered 2 to 3 times along the path – size of the BD and E2E. On the other hand the buffering is applied onto nearly 25% of burst trains hence the average BD of each burst would be 0.062 ms. At the same time this 25% of burst trains require 116 O/E blocks in order to deliver loss-less mode for the traffic of 1.33 Tbps. This traffic is distributed into 7 wavelengths with the maximal efficiency of one wavelength is 86%. Basically, wavelength utilization higher than 30% is very appealing and very difficult to achieve for OCS systems, i.e., any system without all-optical grooming.

Even if the average link utilization would imply considerably higher BBP, BBP is very low for the other tested topologies. Such a low BBP is caused by the combination of multiple factors. First of all, the other networks contain more links and nodes, so the routing algorithm has more options where to route the traffic. Along the topology, more wavelengths are used for those instances. Combination of a more connected topology and more wavelengths allows the routing algorithm to select a path with less aggregation.

Topology		Polska	Nobel US	Germany	Nobel EU
Burst buffering probability	[-]	0.242	0.069	0.068	0.094
Access delay	[ms]	0.022	0.025	0.049	0.061
Buffering delay	[ms]	0.257	0.211	0.252	0.272
End-to-end delay	[ms]	0.642	0.329	0.368	0.476
# wavelengths	[-]	7	9	16	56
Configuration utilization:					
Maximum	[%]	86.12	71.12	73.52	93.13
Average	[%]	68.43	48.49	48.27	44.36
Average link utilization	[%]	72.52	56.84	58.15	56.87
Average # of contenting flows	[%]	92.42	25.64	18.77	28.51
Requested traffic	[Tbps]	1.33	2.18	5.02	13.88
# flows	[-]	121	182	272	2256
Shortest paths	[%]	90.61	85.98	76.39	73.19
Nodes	[-]	12	14	17	28
Links	[-]	18	21	26	41
Diameter	[-]	4	3	6	8
Potential contentions	[-]	116	62	85	579
O/E blocks	[-]	36	60	65	305

Table 6.1: Performance

Another monitored parameter is *Average number of contenting flows* that quantifies the amount of aggregations because the SLE routing avoids primary contention. This parameter should be read as a coefficient specifying how much of traffic is aggregated at a node in the network. At the same time what is left over 100% corresponds to remaining capacity for bypass. Polska network exhibits inclination for aggregation hence the values of BBP and BD correlate to the graphs in Section 6.1.1. On the other hand, the remaining topologies have bypass inclination, thus different BBP and BD values are observed than what was expected based on results in Section 6.1.1.

As one can see, the SLE-RWA algorithm works well with higher amounts of traffic. Higher amounts of traffic allow the algorithm to use more wavelengths. With more wavelengths, the algorithm can efficiently use the SLE. The SLE causes high link utilization and minimal contention. Most importantly this algorithm allows achieving very high wavelength utilization. The main drawback of this algorithm is the weak load balancing as one can see in Figure 6.12. Only a few wavelengths can achieve high wavelength utilization. Generally, the wavelength utilization culminates with increasing wavelength, see Figure 6.12. This behaviour is the direct consequence of traffic assignment to the wavelength configurations where the first wavelength configurations receive majority of traffic. Then all the other wavelength configurations only ensure transmission of the traffic "leftovers" as a sequel they are less efficient.

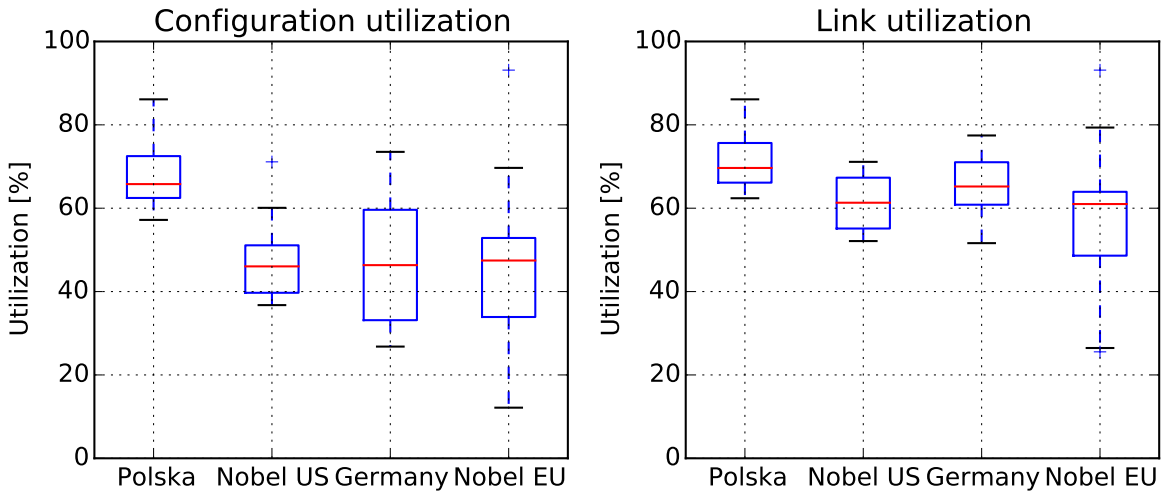


Figure 6.12: Visualisazation of wavelength utilization of tested networks. The red dashed line indicates the average wavelength utilization.

6.3.3 Architecture viability

The simulations exhibit promising time domain parameters, and suggest better load balancing. The CARBOS, however, can not work without O/E blocks that deliver the loss-less properties, and influence the network price. In this section attention was paid to these extra O/E block which are necessary to install onto each CAROBS node in addition to the O/E block principal for the CAROBS node function, i.e., burst disintegration.

The CAROBS network uses WDM, and colourless optical detectors, i.e., O/E blocks, hence the O/E blocks can be shared among some wavelengths. Each O/E block can be shared among all the ingress ports of CAROBS node. The experiment shows O/E blocks requirement correlates with the number of wavelengths. This number amounts the cases when O/E block is necessary with respect to the routing obtained from SLE-RWA. One can see the total number of cases like this in Figure 6.13a), and influence of the Average number of contenting flows parameter that is influenced by the topology complexity. One must see that with more complicated network topology SLE-RWA minimizes the occurrence when secondary contention can happen.

The Figure 6.13b) gives upper limit on the number of O/E blocks per wavelength. The actual number of necessary O/E blocks that are enumerated based on the worst case simulation scenario are captured in Tab. 6.1. One can see that the wavelength and ingress port sharing significantly influences the number of O/E blocks which are necessary for loss-less feature.

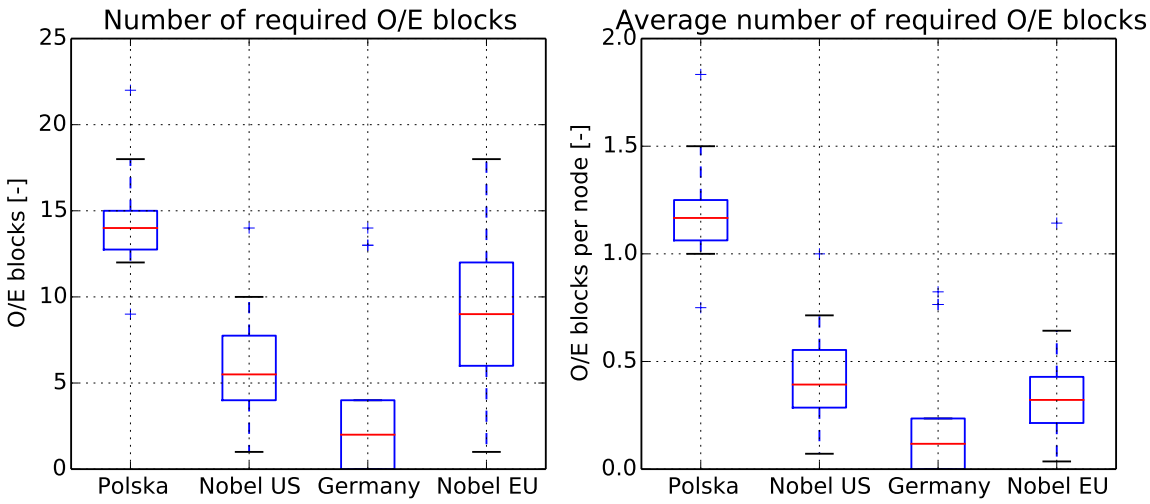


Figure 6.13: Comparison of O/E blocks installations for those four network topologies. a) show average number of O/E blocks per wavelength b) show average number of O/E blocks that must be installed at node for each wavelength (considering colourless O/E blocks).

6.4 Conclusion

This chapter focused on the CAROBS network efficiency through the design of a scalable routing algorithm aiming at exploiting SLE routing. The original motivation od SLE routing was mitigation of contentions because the Phung *et al.* [44] implicates it based on a mathematical model. The CAROBS framework, however, is loss-less, so some contentions occur anyways. With development of loss-less frameworks primary and secondary contentions were distinguish, and SLE routing eliminates only the primary contention. In order to minimize impact of secondary contention an insight into the traffic behaviour is provided, an delaying mechanism limiting the impact of secondary contention is suggested, and the reason why secondary contention can not be fully eliminated is showed. Then the performance of CAROBS node aggregation was showed, but these results only indicate one node behaviour.

The main output of this chapter is a an optimization formulation which is then reformulated into an decomposition formulation providing higher degree of scalability. This decomposition formulation and solution scheme was used in order to find maximal levels of configuration resource utilizations. The configuration utilization translates to wavelength efficiency in real networks hence it is very important to achieve high values because of CAPEX and OPEX savings. The proposed CG SLE-RWA model indicates that the maximal achievable configuration utilization is higher than 70%, which is considerably higher compared with the wavelength utilization of OCS network directly in the optical domain.

This high wavelength utilization is achieved thanks to the grooming on the optical layer, and buffering that provides the CAROBS framework.

Buffering in CAROBS networks means, however, an increase of the E2E delay, for the buffered burst trains are stored in electronic memory for some time. Also there must be installed enough O/E blocks at each node in order to buffer burst trains. If otherwise the loss-less paradigm will never work. The CG SLE-RWA shows locations where secondary contentions can occur and the upper bound on the number of O/E blocks. The accurate number of O/E blocks is obtained using OMNeT++ simulation. Results show big influence of network topology hence analytical formulation is necessary to prepare. Based on other values in the analysis, one can see that the BD is manageable. The E2E can be reduced by limiting the number of buffering along the flow-paths. Nevertheless, the current values without the buffering limit are very reasonable for a network with a link utilization around 70% and more.

The traffic routing distribution, however, needs more investigation because relatively high variance of bandwidth usage among the selected configurations was observed. Partly, this is caused by the traffic pattern where traffic request higher than capacity of wavelength are not expected, but better load balancing is necessary for easier network operation and survivability.

Chapter 7

Software-defined networking and CAROBS networks

The OBS networks are on the verge of being deployed, testing [29] by ISPs has been conducted, a troubling weakness persists in network control. Other authors have focused on OBS performance, but little research on routing information distribution and optical performance monitoring [77] in OBS has been done. This chapters intends to learn from the limitations of transport network control addressed in the field to date and to avoid them in the proposed paradigm. The central control problem in transport networks lies in vendor-specific features that make it difficult to manage devices from vendors in one network, such as GMPLS which offers a wide range of features yet to be provided by all vendors. Considering contemporary trends in network control, the paradigm described in this chapter is inspired from an emerging concept called SDN. The SDN brings an open API to network devices, thereby allowing network resources to be abstracted to high-level applications. This approach simplifies network control and reduces the impact of the vendor lock-in problem by utilizing open control protocols.

The proposed paradigm is based on three building blocks: loss-less OBS for effective transmission, the SDN approach for efficient network control, and distributed computing utilizing external resources (e.g. Cloud computing) to evaluate RWA more easily.

7.1 Software-defined networking

The SDN is a viable concept even in optical networks as shows a survey [78], [79] by Bhaumik *et al.* The main feature of SDN, providing out-of-the box management of forwarding, has been evolving since 1996 [80] as documented by Sezer et al. in [81]. Most recently, the separation of the data plane and control plane was proposed in the OF protocol specification that evolved from the Ethane project published by Casado et al. in [82]

and when subsequent OF development was moved to the Open Networking Foundation (ONF). This consortium continues to improve protocols and has released OF specification in version 1.5 [83]. Although OF is not the only southbound protocol present, it is the most widespread and vendor-supported.

The ONF leads the works on the OF protocol for Ethernet based network. However, the term SDN is not that stable in optical networks. At conferences one can hear different definitions of of optical SDN, and it is everytime dependent to the field of the speaker interest. In the Ethernet based SDN networks, the identification of flows is based on port and different parameters in Ethernet overhead, but in the optical networks, the SDN can have implications to the selection of the modulation format, wavelength, or even transmission technology. Possibilities are almost infinite, and the standardization has not started. Therefore, this chapter can have some impact because the motivation is to deliver very general control framework and compatible with the OF definition.

To our knowledge, only two publications have presented [81], [84] similar research topics. The former paper [84] addresses the problem of the unified control plane and describes GMPLS as being too complicated to be a candidate for the production of an OBS network control plane. The authors of [84] focus on QoS, the signaling of various traffic classes, and a way to inform Edge OBS nodes about appropriate traffic class using the Software Defined Optics interface. The authors of the latter paper [81], which relies on OCS/GMPLS as a control plane, evaluate signaling performance by generating requests at a certain controller load level and verifying response rate. Their experiment is based on an OCS network managed through GMPLS, a NOX controller, and the use of lightpaths for burst transmission. It is highly biased by the OCS/GMPLS network behavior, thus the setup and release times are notably high. The SD-CAROBS concept which is detailed in following sections improves the setup times, completely avoids release time, and increases both overall performance and channel utilization.

7.2 The SD-CAROBS concept

Although the CAROBS framework provides promising performance, see Section 6, it cannot be deployed in its current state as no management and control routines are yet defined. For this reason, a new control plane of the CAROBS framework is propose in this section. Future networks, including CAROBS, should provide a programming interface to provide flexible services based on a customer's immediate needs. Adapting the SDN paradigm is suggested in order to allow convenient network control and easier conversion of current services with optical infrastructure. Consequently, he CAROBS framework was extended to SD-CAROBS based on OF principles. SD-CAROBS is a more general

concept describing the overall architecture, not only the network control mechanism. The four main ideas of the SD-CAROBS are:

1. The SDN Controller and RWA evaluation is distributed and run in a virtual environment. The virtual environment is used to provide higher scalability, reliability, and CAPEX and OPEX savings.

2. One shared wavelength is dedicated to both control messages and the SD-CAROBS header. This control wavelength is split into two sub-channels and is present throughout the SD-CAROBS network.

3. The RWA evaluation is set aside in a separate compute layer. Outputs of the RWA algorithm are used to generate optimal forwarding instructions and traffic classifiers for ingress nodes [49].

4. Forwarding instructions, i.e., a list of particular output ports along the selected flow-path, and traffic classifiers are uploaded only to the SD-CAROBS ingress nodes through control messages.

Generally, the SDN paradigm is based on dispatching multiple control messages that carry forwarding instructions, in the form of matching rules, to a set of network nodes participating in the transmission of a customer's data [81], [84]. Since these messages are encapsulated in Ethernet, they have to be processed at all passing nodes and increasing the total propagation time of each control message. The lightpath is established only when all nodes are configured and acknowledged. This approach, however, cannot be used with all-optical networking (AON) because the propagation speed of an optical burst is considerably higher than the propagation speed of a control message through a control network. This typical SDN approach would result in high burst buffering due to non-existing forwarding instruction, or channel underutilization due to the associated wait in signaling. Therefore, SD-CAROBS has to develop a new way of controlling message distribution. The pre-calculated forwarding instructions are distributed only to network ingress nodes. Each CHP contains all forwarding instructions necessary to signal the flow-path. As a result, there is no need for the lightpath reservation because SD-CAROBS uses an in-advance flow-path reservation schema.

To clarify the used terminology, first terms related to user traffic classification are defined. All traffic incoming to the SD-CAROBS network is considered to be unclassified as long as a matching rule in the ingress node does not exist. In this proposal, the matching rules known from the OF are referred to as classifiers. If the traffic fits a classifier, it is considered to be classified.

7.2.1 General architecture

The SD-CAROBS concept follows the SDN architecture [85], so the SD-CAROBS architecture contains an infrastructure layer, a controller layer. The complete architecture diagram is depicted in Fig. 7.1, and all the layers are described step-by-step from top to bottom in the following text.

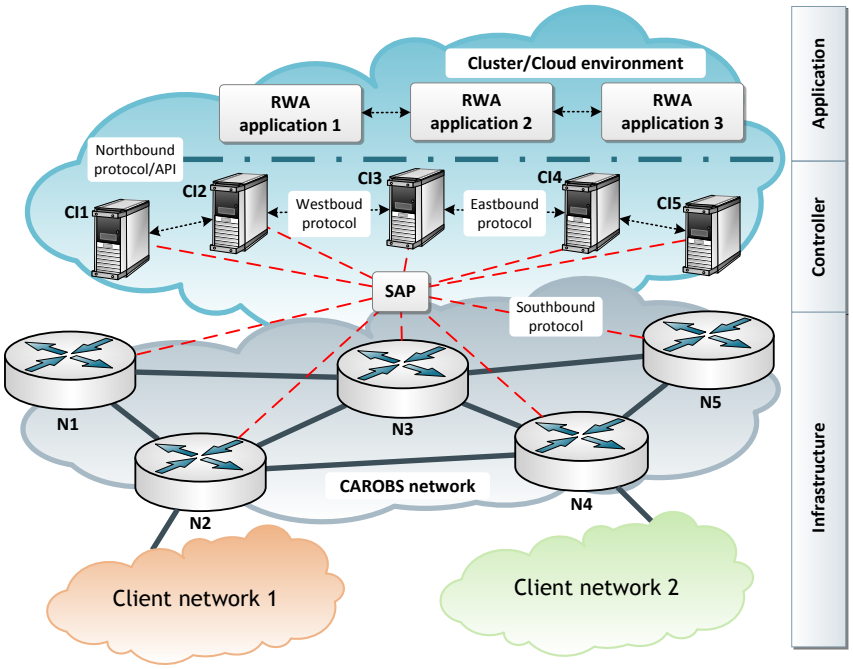


Figure 7.1: High-level view of SD-CAROBS architecture. Both client networks and SD-CAROBS nodes are placed on the base infrastructure layer. Nodes communicate with the controller by southbound protocol, via the SAP. The logically centralized controller, compounded of more controller instances, and the RWA application in the distributed cooperation run on the higher layers. Inter-controller communication is realized via westbound/eastbound protocol, whereas the RWA application accesses the controller layer through a northbound API/protocol.

The application layer sits on the top, and it allows running various control applications, e.g., algorithms to solve the RWA problem in managed network. Running the RWA as an application is vital because it has been proven that RWA is an NP-complete problem [86]. Therefore, solving RWA at each SD-CAROBS node could be performance-demanding and time-consuming. Additionally, calculating the same problem at multiple SD-CAROBS nodes simultaneously would result in a number of duplicities in terms of obtaining the same RWA at multiple SD-CAROBS nodes. More importantly, each node would have to be equipped with high-performance computational resources to calculate a RWA algorithm, and this would increase the network OPEX. The necessity of compute resources could be reduced using heuristics; however, heuristics do not ensure an optimal solution, but heuristics may provide a sufficiently accurate RWA solution in acceptable time. The ISPs

often have an extra computing capacity accessible through the Internet, so this surplus compute capacity can be used when running the RWA application.

The applications, e.g., RWA application, on the compute layer are fed with status data, network statistics, and requests from the controller layer. The result is the global view of the underlying network, RWA application is able to pre-calculate forwarding instructions, wavelength assignment, and traffic classifiers. These descriptors are valid as long as the monitored statistics (e.g. ,number of buffered bursts, offered load) do not exceed preset thresholds. Otherwise, the calculation of RWA application is initiated in order to provide the descriptors that allow achieving efficient RWA configuration of network.

The middle layer contains the logically centralized controller. The controller can be seen as a network operating system or a middleware between applications and network nodes creating infrastructure layer. The interface between the controller layer and the compute layer is defined solely by the controller implementation which usually includes a variety of REST and RESTful API. Conversely, the interface between the controller and the infrastructure layer is limited to a few protocols available currently, e.g., OF. The node-controller communication is ensured by the TLS-secured Transmission Control Protocols (TCPs) protocol on the transport layer. One node is usually connected to multiple instances of one logical controller, or, to multiple independent controllers in traditional SDN architecture.

Even though the availability of the logically centralized controller should be sufficient, the local-closest access to the controller using a SAP is proposed in SD-CAROBS. The SAP is realized via the anycast technique used for various TCPs/IP-based services [87]. Anycast always provides the local-closest connection from a node to the controller to achieve the shortest propagation time of control messages. The shorter the propagation time is, the faster is the user traffic classification. Moreover, the anycast approach decreases configuration complexity because there is just one IP address to connect to.

The SAP represents a logical interface while the real communication among controller instances and SD-CAROBS nodes is realized via the CAROBS control channel. The control channel is realized by a wavelength in CAROBS, but this control wavelength is not fully used, so it is desirable to use FlexGrid technology [59] in order to ensure the SD-CAROBS signaling and node-controller communication. The FlexGrid allows splitting one wavelength into more subchannels and efficient wavelength use for lower data rates, and the most importantly the SD-CAROBS headers can not content with node-controller messages in the time domain. The SD-CAROBS header structure is slightly modified compared to CHP, and the SD-CAROBS header structure is detailed in Section 7.2.2. The control network relies on the IP over Ethernet [88]. This combination is widely used today, and no extra work is required to establish communication in such a network. In

order to avoid loops and broadcast storms, it is beneficial using the spanning tree protocols or its successor, the improved shortest path bridging protocol.

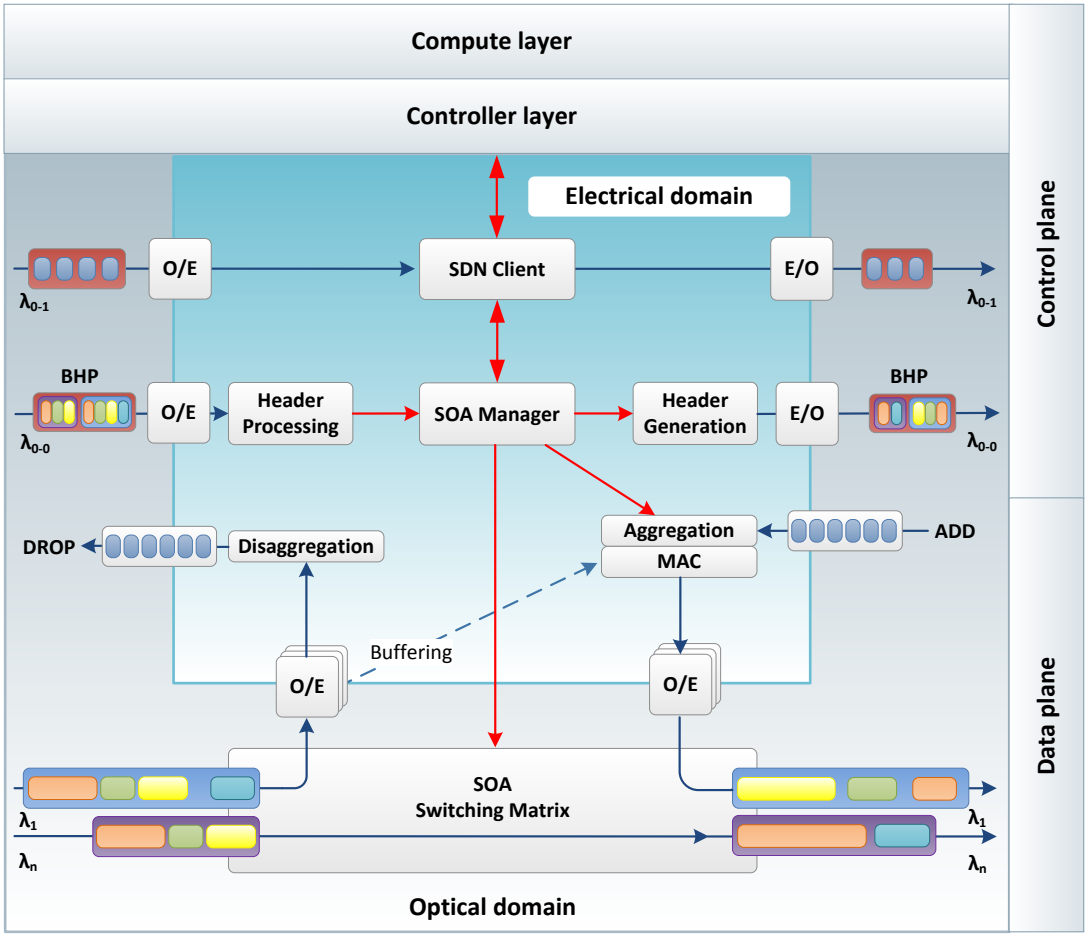


Figure 7.2: The SD-CAROBS node is split into an electrical and optical domain. All control messages and CHPs are converted by O/E and E/O converters. The node's architecture is the same for both Core and Edge nodes. The burst buffering process relies on the O/E block which converts deferred bursts into the node's electrical memory.

The bottom layer represents the underlying physical network, and this is called the infrastructure layer, see Fig. 7.1. The SD-CAROBS nodes operate on the infrastructure layer and they establish connections to the controller layer through the same physical network. The architecture of the SD-CAROBS node is depicted in Fig. 7.2, and the SD-CAROBS architecture is an extension of the original CAROBS node architecture as seen in Fig. 4.10. The new architecture adds an SDN layer represented by the SDN client and dedicated control channel. Each SD-CAROBS node is uniquely addressed according to the OF definition called the DPID. According to the OF specification, the lower 48 bits of DPID correspond to the node's MAC address and the upper 16 bits are left to the implementer. A total of 64 bits gives sufficient address space for an autonomous system run by an ISP.

Topologically, each SD-CAROBS node has a different functionality. The SD-CAROBS

nodes can be located in the core or at the edge of the network. The main purpose of the edge nodes is to aggregate user traffic from adjacent locations. However, the edge nodes can switch burst trains from different edge nodes based on the corresponding SD-CAROBS headers. Even this burst train switching is core node behavior. One can see that unified architecture can reduce network complexity, and also reduces burst contention [49] in the network. Every SD-CAROBS node is controlled from the upper layers as depicted in Fig. 7.2. The SD-CAROBS node is controlled via the SDN client which obtains instructions from the RWA application in the compute layer. Also, the SDN client collects statistical data from the SOA manager and forwards them to the RWA application on the compute layer. This information describe the flows based on the processed SD-CAROBS headers. To deliver all necessary statistical data, currently available OF statistics [89] need to be extended to account for buffered bursts, deferred bursts, memory utilization and buffering delay.

7.2.2 Traffic classification

All the incoming traffic from the client network is distributed to the relevant AQ based on classifiers. When the AQ trigger is initiated, all packets/frames in the AQ are passed to the MAC where these packets/frames are aligned to a burst train, accompanied by a SD-CAROBS header, and sent as part of the burst train to its destination. Once a burst train reaches destination of a particular burst then this burst it is removed from the burst train, i.e., disaggregated. The disaggregation means that the carried user packets/frames are sent to the destination client network using the same method as it was received at the source edge node, e.g., Ethernet.

Every burst transports packets/frames from a triggered AQ, and each AQ is mapped to a set of AP. Each AP maps multiple AQ in order to create a burst train, but each AQ can be the head (first) of only one AP. When the head AQ is triggered, multiple AQs are released and put onto the optical network in the form of a burst train. Bursts composing a burst train are sent along the same flow-path, where the final destination of the flow-path is determined by the first triggered AQ. The identifier of the particular AQ is assigned on the basis of the destination node DPID. The DPID identifier can address more parameters than the destination as defined in the CAROBS framework [49].

The ingress node traffic classification is depicted in Fig. 7.3. In the first step (arrow 1 in Fig. 7.3), the ingress node receives new input traffic from a client network and if this traffic does not fit any classifier, the request for a new classifier is sent to the controller. While the request is pending, the incoming traffic is being buffered into the node's electronic memory. It is imperative that any possible propagation delay of control messages is minimized to avoid memory overflow.

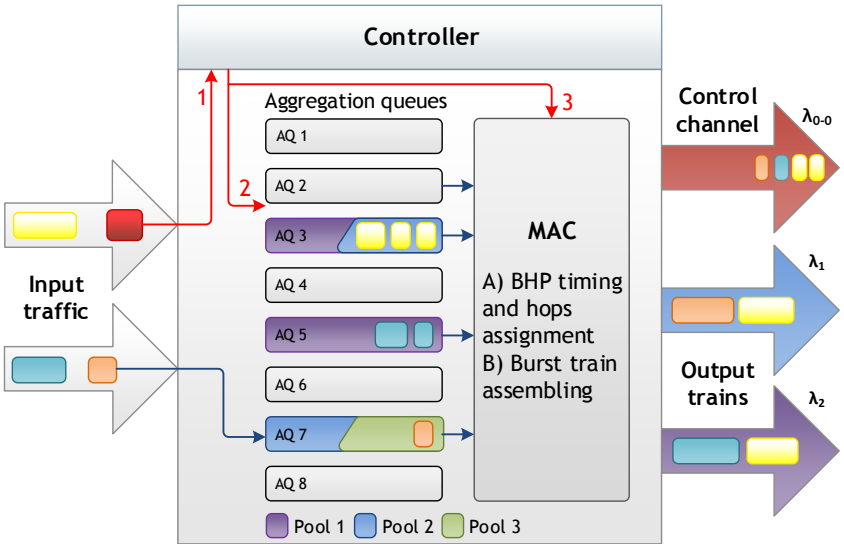


Figure 7.3: Distribution to aggregation queues and train assembling process. 1) New incoming traffic does not fit any classifier so a request is sent to the controller. 2) A new classifier is received and applied. 3) Forwarding instructions, triggers and wavelength are determined by RWA for the corresponding AQ and they are then used to create a SD-CAROBS header.

When the control message is received (arrow 2 in Fig. 7.3) the classifier is stored and traffic is assigned to the appropriate AQ. This message is equivalent to the OF Flow-Mod message [89]. New traffic coming from a client network that hits this classifier is stored in this AQ. The memory space for AQ is allocated dynamically according to the received classifiers. All content of the particular AQ is then emptied into the MAC block in case the AQ is triggered by at least one of two conditions: a time or size trigger.

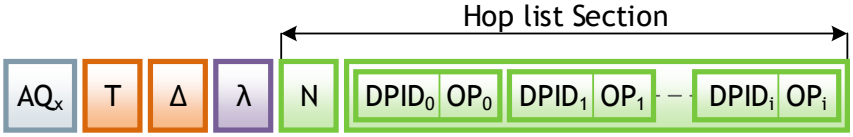


Figure 7.4: Structure of the new Carobs-Flow control message. The Carobs-Flow contains an AQ identifier, time trigger (T), size trigger (Δ), recommended wavelength (λ) and an N dimensional list of DPID/output port number pairs.

The Flow-Mod message is followed by a second control message received in the third step (arrow 3 in Fig. 7.3). This message is called the Carobs-Flow, and its structure is depicted in Fig. 7.4. The Carobs-Flow message contains information about: the AQ identifier destination, time trigger (T), size trigger (Δ), recommended wavelength (λ) and a list of DPID/output port number pairs describing the flow-path. This list has several functions.

Firstly, information in a Carobs-Flow message enables the creation of self-defined APs. The Carobs-Flow message contains a description of the whole flow-path composed

of DPIDs and this list of DPIDs automatically defines a new AP. Every time an AQ is triggered, the content of all AQs that are part of the same AP is assembled to the burst train. This approach significantly reduces the complexity of the internal management of APs.

Secondly, the list of DPID/output ports is used for forwarding in the network. When the lead AQ is triggered, a new SD-CAROBS header is created. This new SD-CAROBS header contains an ordered list of output ports that are used for switching in the network. The SD-CAROBS header is then scheduled for delivery to the designated wavelength λ, via the control channel, to the neighbor node using the first output port from the list of output ports in SD-CAROBS header. The neighboring node checks the SD-CAROBS header and creates instructions for the MX. If any burst from the burst train is not destined for the node, only one forwarding instruction is assigned to MX. Otherwise, two forwarding instructions are sent to MX: one instruction for disaggregation and the second one for forwarding between two ports. The output port for both cases is obtained from the SD-CAROBS header's output port list. If the output port is already busy, the burst train is redirected to the internal electric buffer as depicted in Fig. 7.2. Each node applies a head drop action to the hop list. This action keeps forwarding instructions for the following nodes at the first list position so there is no need for a list pointer.

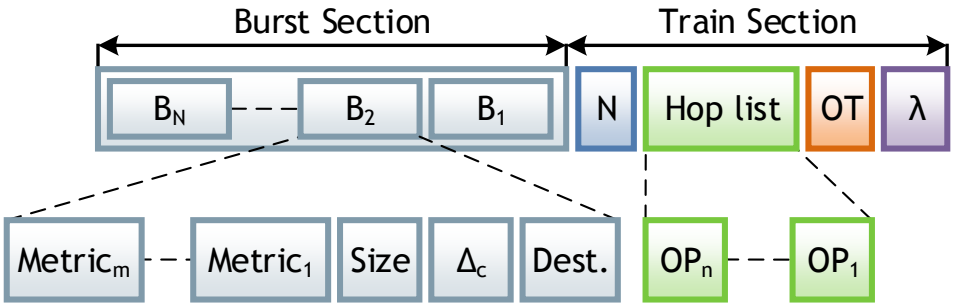


Figure 7.5: CHP structure was extended replacing the Destination field from the original CAROBS design with a new Hop list. The Hop list (representing forwarding instructions) is comprised of output port numbers of every node pertaining to the flow-path in descending order. The Burst Section holds control information of particular bursts waiting for the designated wavelength.

7.2.3 Analysis of control delay

Previously published concepts do not perform well in terms of flow-path establishment. These concepts signal the flow-path as if it was a lightpath, i.e., the wavelength has to be reserved before transmission begins. The lightpath establishment is time-consuming as the reservation must be acknowledged by all nodes. A considerable improvement of set up times in SD-CAROBS comes from in-advance signaling using the Carobs-Flow message.

Additionally, the lightpath establishment approach reduces wavelength efficiency, does not allow all-optical grooming, and can result in channel underutilization due to the delay in waiting for acknowledgment.

Therefore, the propagation delay of SD-CAROBS control messages in this sub-section needs to be analyzed. It can not be omitted that the only significant delay in SD-CAROBS is introduced when new unclassified traffic reaches the ingress node. In this situation, the first packet is sent to the control plane through the SAP. Since the SAP provides a local-closest connection to the controller, it is assumed that a RWA solution is calculated and made available at the controller instance running in a cloud service in the nearest data center. The signaling performance can be verified by a best- and worst-case scenario. In the best-case scenario, the SAP is just one hop away. In the worst case scenario, the nearest SAP interface is on the opposite side of the network. The propagation delay can be quantified by formula (7.1):

$$\mathbf{D_{propag}} = D_{SRC} + 2\left(\sum_{i \in V_S} \ell_i + \sum_{n \in N_S} D_{p,n}\right) + D_{SAP}, \tag{7.1}$$

where index S represents a set of nodes along the shortest path between the SAP and the ingress node. N represents a set of nodes and V denotes a set of vertices in graph G representing a real network. D_{SAP} stands for the shortest processing delay that is required to evaluate a requested flow-path and classifiers on both compute and controller layers. In a further analysis D_{SAP} is 10 ms. D_{SRC} is the processing delay of the source ingress node requesting classifiers from distributed controller, and it is assumed to be 400 μs [90]. The $D_{p,n}$ represents control message processing time. Time depends on the switching performance of the node n, and it is assumed $D_{p,n}$ is 100 μs.

To obtain propagation delays, network topology of Poland, the United States, and Germany [76] were evaluated. The real distances between nodes was calculated using the well-known haversine formula. Since SDN allows to abstract the network one can see its topology as a weighted graph. The best case is a link with the lowest weight and the worst case is quantified using the weighted network diameter. Propagation delays are captured in Table 7.1.

Table 7.1: Control message propagation delay analysis.

	Polska	**Nobel US**	**Nobel Ger.**
$d(G)$ [km]	810.35	4453.15	789.75
$\bar{\ell}$ [km]	187.96	1086.55	143.24
D_{worst} [ms]	19.13	54.58	19.33
D_{best} [ms]	11.37	13.48	10.89

SD-CAROBS offers considerable improvement of control message propagation delay compared to the OCS/GMPLS-based approach [91]. This improvement is not negligible even in the worst-case scenario. The most significant feature is that the propagation delay is not introduced for every flow, but only once for new, unclassified traffic. In other words, SD-CAROBS occasionally introduces a propagation delay rather than every time a transmission is started.

Additionally, compared the SDN approach to the classical MPLS-controlled networks, there is no need for any other inter-node signaling or label propagation protocol. The absence of inter-node signaling makes the SD-CAROBS network easier to maintain, and the only additional communication is the upstream of statistical and status data sourced from the SDN client to the controller through SAP.

7.3 The SD-CAROBS performance

The signal propagation speed is essential for the transmission system. However, without an efficient transmission mechanism SD-CAROBS would not be viable. SD-CAROBS transmission performance depends on the distribution of traffic in the network. Therefore, to distribute traffic optimally, a new a routing algorithm is defined in Section 7.3.1. The routing algorithm is further verified in terms of latency and overall network performance in Section 7.4.

7.3.1 Traffic routing

Nowadays, traditional protocols (Open Shortest Path First (OSPF), etc.) are used in IP networks. These protocols establish forwarding through communication with routers in the same network. Based on this communication, each router establishes its own forwarding policy. These protocols can be easily used as long as the forwarding decisions are based on topology only. When the forwarding decision takes into account the traffic flows within a network, calculating forwarding can last a long time even for relatively small testbeds sizes [92]. Multiple wavelengths are used in optical networks, so the problem of finding forwarding is even more demanding. The forwarding decision follows similar rules as in optical networks, but is calculated from multiple wavelengths at the same time. Consequently, one can see the routing problem as traffic routing on multiple topologies which are equal, i.e., virtual topologies. Therefore, to obtain an optimized traffic distribution, traditional protocols based on shortest paths cannot be used hence more advanced techniques must be considered.

As our aim is a high performance with SD-CAROBS, an accurate RWA model using traffic grooming is defined. The traffic forwarding application calculating this RWA can be

found on the application layer of SD-CAROBS. RWA as a mathematical model is formulated in order to obtain routing information spanning a minimal number of wavelengths. The formulation is a mathematical model which selects the best wavelength configurations, where a wavelength configuration consists in a set of provisioned flow-paths on a given wavelength. Consequently, it has been discovered that burst loss increases as increases the number of flows merging into one egress from multiple ingress ports. This observation was named the streamline effect [70]. The RWA model relies on variables which are defined below:

$\mathcal{SD}$ – set of node pairs such that $D_{sd} > 0$
D_{sd} – demand for node pair (v_s, v_d)
B – wavelength transport capacity
C – set of wavelength configurations, indexed by c
c – generic wavelength transmission: it is characterized by a set of paths p^c so that no paths merge (thanks to the streamline effect). Each path is characterized by

- $p_\ell^c = 1$ if the link ℓ belongs to the path, 0, otherwise,
- $b_\ell^c =$ amount of bandwidth routed on ℓ in configuration c,
- $b_{sd}^c =$ amount of bandwidth routed in c for sd,

z_c – positive integer variable $z_c \in \mathrm{Z}^+$ specifying for how many wavelengths, wavelength configuration c will be used.

The RWA model objective function:

$$\min \sum_{c \in C} \mathrm{COST}_c \, z_c \tag{7.2}$$

which says that minimum number of wavelength configurations will be used. COST_c represents the quality of the wavelength configuration. The direct implication is that only good wavelength configurations are selected, aiming at using a minimum number of wavelengths. The objective function (7.2) is subject to:

$$\sum_{c \in C} d_{sd}^c z_c \geq D_{sd} \qquad \{v_s, v_d\} \in \mathcal{SD} \tag{7.3}$$

$$z_c \in \mathrm{Z}^+ \qquad c \in C. \tag{7.4}$$

Constraints (7.3) correspond to the demand constraints.

7.3.2 RWA algorithm solution

The selection of wavelength configurations defines the MP. If restricted to a selection among a subset of all potential configurations, it defines the so-called RMP in terms of OR. RMP is initially defined by a set of simple, pre-calculated wavelength configurations that are not necessarily part of the optimal set of wavelength configurations. Next is define the PP, which can generate an improving wavelength configuration, i.e., a configuration that, if added to the current RMP, can improve the optimal value of its linear relaxation. The PP, an ILP, generates a new wavelength configuration, with an objective function corresponding to the reduced cost [32] of variable z_c. If the the value of the reduced cost is negative, the wavelength configuration will improve the optimal value (linear relaxation) of the current RMP, otherwise it means more wavelength configurations are note necessary to generate as the optimal solution of the linear relaxation of the MP have been reached. This process is depicted in Fig. 6.10 as *LP Solution of MP*. Once the optimal solution of the linear relaxation is reached, it remains to derive an integer solution, either using an exact algorithm (e.g., branch-and-cut method) or a heuristic, i.e., the optimal integer solution of the last RMP.

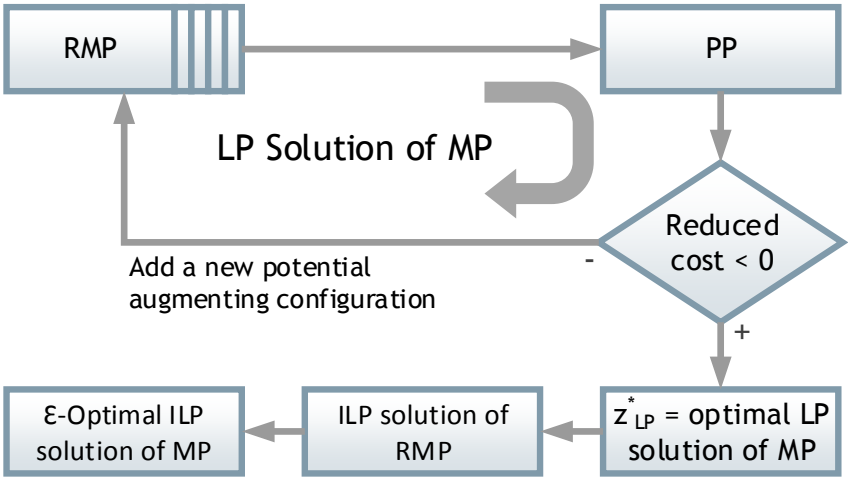


Figure 7.6: Flowchart of the solution process.

OBS network performance is limited by contentions caused by bursts coming from different ingress ports and heading to the same egress port [44] on the same wavelength. Therefore, to minimize the impact of the contention, wavelength configurations that limit flow merging as much as possible are generated, and PP consists of constraints that *a)* find flow-paths in a graph, and *b)* count the number of merging flows. The number of merging flows is then used as the parameter defining quality of the wavelength configuration COST. The PP is defined as a optimization model which relies on the variables defined below:
b_{sd} – amount of requested bandwidth in the configuration from v_s to v_d,
$\varphi_\ell^{sd} \in \mathbb{R}^+$ – provisioned bandwidth in the configuration from v_s to v_d on link ℓ,

y_ℓ^{sd} – decision variable such that $y_{sd}^\ell = 1$ if some traffic from v_s to v_d is routed on link ℓ in the configuration,
$y_{\ell_i \ell_o}$ – decision variable indicating that sd flow is routed from ℓ_i (incoming port) to ℓ_o (outgoing port) at node $v = \text{SRC}(\ell_o)$.

Objective function (reduced cost):

$$\overline{\text{COST}} = \text{COST} - \sum_{\{v_s, v_d\} \in \mathcal{SD}} u_{sd}\, b_{sd} \tag{7.5}$$

Constraints defining a wavelength configuration:

$$\sum_{\ell \in \omega^+(v)} \varphi_\ell^{sd} - \sum_{\ell \in \omega^-(v)} \varphi_\ell^{sd} = \begin{cases} b_{sd} & \text{if } v = v_s \\ -b_{sd} & \text{if } v = v_d \\ 0 & \text{otherwise} \end{cases} \tag{7.6}$$
$$v \in V, \{v_s, v_d\} \in \mathcal{SD}.$$

This first set of constraints defines routes from v_s to v_d. Routes do not necessarily correspond to the shortest-paths in the network as the goal is to minimize the number of wavelengths, while minimizing the number of merging flows. If the shortest-path approach was used it would result in the use of more wavelengths.

In this routing algorithm split source routing was not allowed; therefore, multiple flows were not allowed to start at node s by (7.9) and (7.10). Additionally, a sd flow can pass through an intermediate node only once (7.11). In other words, these constraints blocks loops that might be generated by (7.6).

$$\varphi_\ell^{sd} \le D_{sd} y_\ell^{sd};\ y_\ell^{sd} \le \varphi_\ell^{sd} - \varepsilon \qquad \ell \in L, \{v_s, v_d\} \in \mathcal{SD}, \quad \varepsilon \to 1 \tag{7.7}$$

$$\sum_{(v_s, v_d) \in \mathcal{SD}} \varphi_\ell^{sd} \le B \qquad \ell \in L \tag{7.8}$$

$$\sum_{\ell \in \omega^-(v_s)} y_\ell^{sd} = \sum_{\ell \in \omega^+(v_d)} y_\ell^{sd} = 0 \qquad (v_s, v_d) \in \mathcal{SD} \tag{7.9}$$

$$\sum_{\ell \in \omega^-(v_d)} y_\ell^{sd} \le 1;\ \sum_{\ell \in \omega^+(v_s)} y_\ell^{sd} \le 1; \qquad \{v_s, v_d\} \in \mathcal{SD} \tag{7.10}$$

$$\sum_{\ell \in \omega^-(v)} y_\ell^{sd} = \sum_{\ell \in \omega^+(v)} y_\ell^{sd} \le 1 \qquad v \in V : v \notin \{v_s, v_d\}, \quad (v_s, v_d) \in \mathcal{SD} \tag{7.11}$$

Constraints (7.8) allow multiple flows to be routed onto one wavelength provided they do not exceed the transport capacity given by the wavelength data rate. This formulation is enabled by OBS allowing traffic mixing due to sub-wavelength channel scheduling which is allowed by JET or JIT signaling protocols in OBS networks, and by JET in SD-CAROBS networks.

$$y_{\ell_i}^{sd} + y_{\ell_o}^{sd} - 1 \leq y_{\ell_i \ell_o} \qquad \ell_i, \ell_o \in L : \text{DST}(\ell_i) = \text{SRC}(\ell_o) = \{v\}, \\ v \in V, (v_s, v_d) \in \mathcal{SD} : v \notin \{v_s, v_d\} \tag{7.12}$$

$$\sum_{\ell_i \in L : \text{DST}(\ell_i) = \text{SRC}(\ell_o)} y_{\ell_i \ell_o} = \text{COST} \qquad \ell_o \in L \tag{7.13}$$

$$y_\ell^{sd} \in \{0, 1\} \qquad \ell \in L, \{s, d\} \in \mathcal{SD} \tag{7.14}$$

$$d_{sd} \geq 0 \qquad \{s, d\} \in \mathcal{SD} \tag{7.15}$$

$$\varphi_\ell^{sd} \geq 0 \qquad \{s, d\} \in \mathcal{SD}, \ell \in L \tag{7.16}$$

$$y_{\ell_i \ell_o} \in \{0, 1\} \qquad \ell_i, \ell_o \in L : \text{DST}(\ell_i) = \text{SRC}(\ell_o). \tag{7.17}$$

As stated, contentions are caused by merging flows in OBS networks, therefore, constraints evaluating the number of merging flows were introduced. The objective was to minimize the total number of merging flows in network. Constraints (7.12) identify flows that are switched from the ingress port connected to link ℓ_i of node v to the egress port connected to link ℓ_o. Variable $y_{\ell_i \ell_o}$ is equal to one if there is at least one flow $\ell_i \rightarrow \ell_o$, zero otherwise. Constraints (7.13) compute the number of merging flows for each wavelength configuration. This total number of merging flows is addressed as COST and used as the parameter quantifying the quality of the wavelength configuration in RMP.

7.3.3 Wavelength configuration assignment

When the *LP Solution of MP* loop stops an an integer solution, it means a solution has been derived then the resulting selected set of wavelength configurations corresponds to an optimized combination of wavelength configurations.

The accuracy of the RWA solution as derived by the algorithm described in the flowchart of Figure 6.10 can be then assessed by

$$\epsilon = \frac{\tilde{z}_{\text{ILP}} - z_{\text{LP}}^\star}{z_{\text{LP}}^\star},$$

where $z_{\text{LP}}^\star$ and $\tilde{z}_{\text{ILP}}$ denote the optimal LP and optimized ILP values of the RWA solution.

7.4 Performance analysis

The SD-CAROBS performance analysis represents real network behavior that can be seen by both providers and users. So far optimal traffic distribution was obtained in the network from RWA described in Section 7.3.1. The traffic routed according to this distribution is affected by sub-wavelength scheduling in the time domain. Therefore, the routing distribution fails to quantify SD-CAROBS performance. To quantify SD-CAROBS performance, a time domain analysis, i.e., simulation, was carried out. Our focus was on the E2E, access delays and statistical parameters of SD-CAROBS, such as buffering probability. Each number represents a mean value over all measured data from each node. Similarly, confidence intervals are not provided for these values because the standard error values are negligible compared to the mean values. The simulations were conducted using an OMNeT++ simulator and modified CAROBS models [53]. Source nodes s supplied traffic generated according to the Poisson distribution. The generated payload packets of constant size (100 kb) representing the flow were sent to aggregation queues to generate bursts. JET [13] was used as a signaling protocol and LAUC-VF algorithm [68] for burst assembly.

The topologies used for the SD-CAROBS performance analysis are depicted in Fig. 7.7 and the topological information was used from SNDlib [76]. However, since the traffic D_{sd} provided by SNDlib was not sufficient, D_{sd} was randomly generated using a uniform distribution.

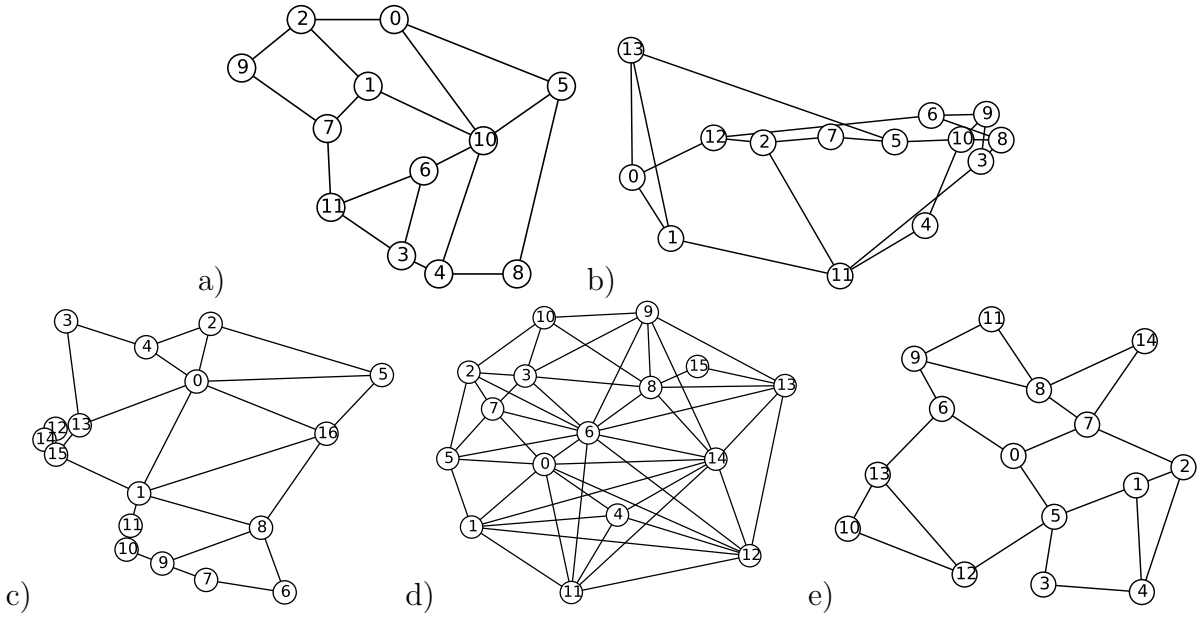


Figure 7.7: Network topologies used for analysis: a) Polska b) Nobel-US c) Germany d) New York e) Atlanta, originated in SNDlib [76]. The a)-c) reflect real network and d)-e) reflect artificial networks often used by researchers in their works.

The results obtained for these topologies are captured in Table 7.2 which contains three sections a) topological information where each topology is quantified from the viewpoint of network complexity; b) routing performance which quantifies the traffic distribution for given traffic patterns obtained from RWA algorithm described in Section 7.3.1; c) the SD-CAROBS performance analysis where information about the time domain performance were capture – simulations. Part c) is noteworthy as it provides performance information that can be measured in a real network. However, it is necessary to consider that every model relies on a number of approximations, thus a gap between real measurements and simulations is seen. In our case, the most crucial approximations relate to input traffic characteristics. In this analysis, the packets from the connected network are assumed to arrive following a Poisson distribution at a rate of α packets per second so that the inter-arrival time between packets is a negative exponential distribution with parameter α. SD-CAROBS uses both time and space triggers, therefore, burst assembly tends toward Gaussian distribution asymptotically [40], according to the central limit theorem [59]. In other words, the impact of the input traffic distribution is negligible due to the traffic encapsulating into bursts.

It is worth noting that the wavelength utilization achieved by SD-CAROBS is more than 80%, which is significant when compared with OCS RWA [93]. It is generally known, that OCS networks are used with up to 40% of wavelength efficiency. This tremendous SD-CAROBS wavelength utilization was achieved by the statistical multiplexing which OBS provides: a link and wavelength are used by a number of flows concurrently. The price paid for the statistical multiplex is represented by the burst loss which results to the BBP for SD-CAROBS. Needless to say, the BBP is slight. One can see that as the wavelength data rate increases the, BBP become smaller. In simulations the same type of modulation for all data rate scenarios were assumed. Meaning, time scheduling works in the same manner; bursts have approximately the same length, but the traffic from adjacent networks is higher. If a more powerful modulation formats were used shorter bursts would be obtained. These shorter bursts are better scheduled, so even lower BBP can be observed.

The time domain meters show that buffering in the SD-CAROBS network can significantly increase the delay of the whole burst train. Based on the BD and the E2E, one can see that bursts are buffered more than once along the flow-path. This means the E2E can be reduced by modifying the RWA algorithm. However, it may result in lower wavelength utilization.

Either way, the SD-CAROBS shows very appealing performance parameters. These parameters are archived by grooming in time domain which causes extra contentions that are mitigated by electrical buffering. This buffering is enabled by the O/E blocks, and

Table 7.2: Results of performance analysis, topological information, and traffic routing quantification.

Topology		Polska	Nobel US	Nobel Germany	New York	Atlanta
Nodes	[-]	12	14	17	16	15
Links	[-]	18	21	26	49	22
Diameter	[-]	4	3	6	3	5
Category		Real	Real	Real	Theor.	Theor.
Routing performance						
Max. wavelength utilization	[%]	84.31	87.54	90.94	88.45	90.79
Average wavelength utilization	[%]	62.28	73.13	65.91	62.91	68.93
Max. link utilization	[%]	84.31	87.54	90.94	88.45	90.79
Average link utilization	[%]	76.55	79.26	87.22	82.22	89.66
Number of wavelengths	[-]	10	12	23	8	19
Traffic performance at 10 Gbps						
Routed traffic	[Tbps]	2.13	3.51	7.47	4.70	5.50
Burst buffering probability	[-]	0.228	0.304	0.238	0.268	0.261
Access delay	[ms]	0.006	0.115	0.100	0.073	0.100
Buffering delay	[ms]	0.254	0.285	0.294	0.276	0.276
End-to-end delay	[ms]	0.589	1.082	1.041	0.929	0.991
Essential O/E blocks	[-]	101	162	252	247	214
O/E blocks for loss-less	[-]	58	117	213	131	174
O/E efficiency margin	[-]	0.392	0.262	0.142	0.166	0.178
Traffic performance at 100 Gbps						
Routed traffic	[Tbps]	21.3	35.1	74.7	47.0	55.0
Burst buffering probability	[-]	0.189	0.257	0.177	0.224	0.200
Access delay	[ms]	0.024	0.056	0.046	0.080	0.063
Buffering delay	[ms]	0.652	0.632	0.685	0.523	0.676
End-to-end delay	[ms]	0.687	2.002	1.436	0.602	1.280
Essential O/E blocks	[-]	109	173	268	254	217
O/E blocks for loss-less	[-]	46	101	182	114	145
O/E efficiency margin	[-]	0.402	0.267	0.146	0.171	0.190
Traffic performance at 200 Gbps						
Routed traffic	[Tbps]	42.6	70.2	149.4	94.0	110.0
Burst buffering probability	[-]	0.158	0.224	0.146	0.193	0.168
Access delay	[ms]	0.021	0.067	0.050	0.018	0.074
Buffering delay	[ms]	0.533	0.833	0.829	0.684	0.912
End-to-end delay	[ms]	0.810	2.273	1.285	1.515	2.595
Essential O/E blocks	[-]	101	170	235	253	224
O/E blocks for loss-less	[-]	37	89	203	101	127
O/E efficiency margin	[-]	0.282	0.451	0.150	0.178	0.196

Table 7.3: Quantification of link-disjoint routing realizing RWA of an OCS network.

Topology		Polska	Nobel US	Nobel Germany	New York	Atlanta
Max. wavelength utilization	[%]	72.86	71.15	77.11	66.82	71.71
Average wavelength utilization	[%]	35.97	36.24	36.47	29.08	34.82
Max. link utilization	[%]	74.97	72.89	80.20	69.87	76.73
Average link utilization	[%]	35.97	36.24	51.07	47.80	52.29
Number of wavelengths	[-]	12	13	23	8	20
Number of flows	[-]	143	209	322	262	237
O/E efficiency margin	[-]	0,252	0,173	0,159	0,182	0,221

directly influence the node and network CAPEX and OPEX. Therefore, the comparison of SD-CAROBS and OCS is created. In order to compare SD-CAROBS with OCS the

pricing problem is modified such that the (7.13) constraints and the COST are removed from (6.22). Then a new set of constraints allowing generation of link-disjoint flow-paths, i.e., lightpaths, was added. These constraints are formulated as (7.18).

$$\sum_{\{s,d\}\in\mathcal{SD}} y_\ell^{sd} \leq 1 \qquad\qquad \ell \in L \tag{7.18}$$

The RWA generating link-disjoint lightpaths and used traffic pattern requires more wavelengths in order to support the same traffic and gives worse values for average utilizations of wavelength and links, see Table. 7.3. The parameter *number of flows* implicates the number of O/E blocks that are necessary in OCS networks. The number of flows is higher than the number combination of destinations $|V|.(|V|-1)$ because some flows are split among multiple lightpaths. In OCS networks lightpaths are established for long term then the O/E block sharing is not possible, for each O/E block is exclusively dedicated to each lightpath. On the other hand, each O/E block can be shared by multiple flows in SD-CAROBS networks thus the number of *essential O/E blocks* is considerably smaller than number of O/E blocks necessary in OCS RWA. The total number of O/E blocks, however, is higher because there are some additional O/E blocks necessary in order to mitigate burst loss. The absolute numbers can be misleading thus the contribution of each O/E block is quantified by *O/E efficiency margin* parameter which tells how much of wavelength efficiency each O/E block contributes. The higher value of this parameter indicates better efficiency of each O/E block. Table 7.2 shows that Polska and Nobel US topologies are very efficient using O/E blocks, but the others are less efficient than the OCS. This can be seen as a price for exceeding the 40% wavelength efficiency of OCS.

7.5 Conclusion

The SDN approach for network control was adapted as SDN readily allows integration with other network function blocks in ISP networks. SDN is an promising network control approach that some ISPs employ in their networks [94] today as new services can be deployed, e.g., bandwidth on demand, or a virtual network. For transmissions on the optical layer, this concept proposes to use the CAROBS framework to ensure loss-less transmission and grooming directly on the optical layer to provide high network efficiency. This can only be achieved through an innovative approach of flow-path establishment, and efficient burst train forwarding based on logically selected flow-paths calculated by the SD-CAROBS controller. These paths can be calculated by the SD-CAROBS controller thanks to statistical traffic information collected in the network.

The performance analysis shows the maximal achievable performance of the SD-CAROBS network when the routing minimizing merging is used. An RWA algorithm was formulated to produce the routing which was used in an OMNeT++ simulator in order to obtain traffic performance parameters. The obtained traffic performance parameters indicate highly satisfactory values for processing delays in the network. This is compelling as the wavelength configurations achieve more than 84% utilization directly on the optical layer. Such a high-wavelength efficiency significantly outperforms any OCS-based system, such as the optical transport hierarchy or Ethernet over WDM.

Chapter 8

Conclusion

8.1 Summary of thesis

The viability of OBS networks is still a question, so I focused on the viability of the next generation iteration, i.e., loss-less OBS paradigm. Specifically, I worked with the loss-less OBS concept called CAROBS framework, and I further extended it by a new network control paradigm. I started the research with detail analysis of OBS networks, and this analysis led to the implementation of simulation models of the loss-less OBS framework CAROBS. Using these simulations models implemented in OMNeT++ event driven simulator, I carried out first experiment dealing with performance analysis of CAROBS networks [53], and I showed that deeper research is necessary.

Subsequently, I did an indepth analysis of the CAROBS network traffic behaviour, i.e., loss-less OBS network viability analysis. Results gave recommendation for using colourless detectors and WDM systems with more than 10 wavelengths. This conclusion is compatible with the trend of optical networks today. Subsequently, the analysis allowed me to analytically formulate the CAPEX and OPEX. However, this formulation was very complex, so I could not use it for optimizations using the approach of mathematical programming, i.e., LP.

Consequently, I selected a corner case of this analytical formula. This corner case is based on the SLE approach, and it allowed to formulate GRWA using techniques of mathematical modelling. This formulation, however, is still too complex to be viable for bigger networks due to scalability, so I decided to use decomposition techniques in order to take advantage of the mathematical model structure. Specifically, I selected CG approach then I created the CG reformulation of the original ILP formulation. The CG formulation is less accurate, but still provides optimal solution compared to some heuristics, i.e., still provides the best possible solution. This solution shows very appealing wavelength efficiency in mesh AON hence I showed that loss-less OBS architectures are viable.

This CG SLE-RWA formulations is still comprehensive, so traditional approach of its calculation is not possible to use. Nevertheless, calculation of SLE-RWA is pointless because SLE-RWA calculates routing strategy for whole network not for a particular node hence multiple same results would be gotten. Consequently, centralized control paradigm is designed. This proposed paradigm is called SD-CAROBS because it is based on the SDN approach for network control. Currently, the SDN is seen as future for the converged networks hence it is very good to define compatible network control because then it is easier to implement such an AON. Defining the SD-CAROBS, it is possible to control optical infrastructure and environement for services from one place hence more efficient use of all resources in ISP network can be achieved. In other words, SD-CAROBS allows to decrease ISP's CAPEX and OPEX.

8.2 Fulfilments of targets

All the goals that were set in Section 2 were **achieved**. Each goal's fulfillment is commented bellow:

1. **Analysis of loss-less OBS traffic properties, and definition of the working conditions under which the selected loss-less OBS framework can be used. Discussion the impact of loss-less OBS traffic properties onto the loss-less OBS viability.**

   The discussion of the CAROBS framework behaviour is provided in Section 5. In order to obtain significant number of replications, I used the OMNeT++ models and computer clusters[1], so the results are statistically valid. Additionally, in order to find the CAROBS node stability threshold, I used the MVA analysis. Then I found the stability threshold, and the its dependence on the number of merging flows.

   Most importantly I quantified the dependence of the number of O/E blocks, so I showed the relation between the number of O/E blocks, the node load, number of merging flows, and number of wavelengths that implicates the CAROBS node price which results in CAPEX and OPEX of whole CAROBS network.

2. **Address the viability of loss-less OBS architecture and optimizations of loss-less OBS networks.**

   I synthesized the results of the previous target in the Section 5.5 and created the pillar for the formulation describing relation among the node load, number of wave-

[1]Used clusters: Metacentrum, Briaree, Mammouth parallèle 2, Mammouth série 2, and Guillimin that are part of Calcul Quebec. Simulations took around 13.000 Central Processing Unit (CPU) per day.

lengths, and number of merging flows such that CAPEX and OPEX of a CAROBS network is possible to quantify.

The most importantly, this formulation is formulated as an ILP program. However, this ILP program contains some non-linear constraints, so it is not possible to solve it using an solver, such as Cplex [37]. Therefore, the heuristics must be used in order to use this formulation, but the heuristics can be very inaccurate compared to ILP.

3. **Formulate the optimization mechanism of the selected loss-less OBS architecture and show the maximal performance of it.**

   The previous target showed the direction to follow in order to obtain an affordable CAROBS network. Therefore, I focused on the special case called SLE in order to exploit the CAROBS network performance. The SLE gives to create the RWA algorithm called SLE-RWA, and this RWA algorithm is described in the Section 6.

   In order to find the maximal performance, I modified this ILP to the form of a decomposition model because of the ILP model scalability issue. I decomposed the ILP into two ILP models which are closely related and together mimic the original ILP model, but scales. This model shows very high wavelength efficiency for all studied network topologies. Compared to well know OCS wavelength efficiencies, i.e., bellow 40%, the SLE-RWA allows achieving wavelength efficiency around 80%.

4. **Define a control mechanism which is usable for OBS architecture and compatible with current technology. The proposed control mechanism must easily integrate into currently maintained networks. Also, this control mechanism must allow delivering the maximal network performance**

   The current trend of network control tends to SDN which brings many appealing features, and the SDN tackles network traffic as flows which are very similar to flow-paths used for the CAROBS network traffic quantification. Therefore, it is very convenient to use SDN paradigm for CAROBS network control. I described the SDN aware CAROBS network control in Section 7.2.

   The SD-CAROBS decouples RWA from node into a distributed environment, e.g., cloud services. Therefore, SLE-RWA can be used for RWA, so very high network, wavelength, efficiency can be achieved.

5. **Implement simulations model using an event driven simulator in order to have a tool for verification of studies of the loss-less OBS networks.**

   I discussed the loss-less OBS architectures in the section dealing with OBS network architectures. I wrote this section in order to introduce a reader into the field of

all-optical networks using the OBS paradigm. The discussion of loss-less paradigms is provided in Section 4.2.

Based on the discussion of loss-less architectures the CAROBS architecture was selected. The simulation models were implemented using OMNeT++. OMNeT++ is becoming the standard because it is OpenSource and allows very versatile use, see Section 3.1.1. The implemented OMNeT++ models of CAROBS were first time published in [53].

8.3 Further extensibility and recommendations

The contribution of this work is in two main areas. It is the traffic RWA and network control. Currently, results of both areas only show direction, but plenty of work still needs to be done. However, the research in OBS network culminates, and it is questionable whether sub-wavelength granularity achieved by traffic grooming in time domain is really viable or the future direction in all optical networks is grooming based on FlexGrid. Either way SLE-RWA is still applicable for FlexGrid. The FlexGrid is very promising direction in AON hence it is beneficial to pay some interest to it. My recommendation for future research and extension of my work are as follow:

- Study the just in time control mechanisms in FlexGrid networks in order to deliver an AON able of fast reconfiguration and fine network resource allocation.
- Study the exploitation of SLE-RWA for FlexGrid networks and verify SLE-RWA viability in FlexGrid networks.
- Verify the viability of the SD-CAROBS for traffic control in datacenter in order to deliver a very fast switching platform which easily allows offloading of jumbo flows.
- Study the offloading using SD-CAROBS among datacenters in case of a failure in the datacenter or whole datacenter in order to minimize service, that runs in the affected datacenters, downtime.
- Study the server deployment problem among datacenters in order to create a survivable and sustainable service in case of datacenter failure.

Appendix A

Universal data analyser

The UDA is a library developed in order to analyze OMNeT++ simulations and carry out complex analysis. The work on very first version of this library started in the year 2012, and this basic library used plain Python for mathematical evaluation and Matplotlib [95] for visualization. This version was not named UDA and OMNeT++ was the only input format. Later, other input formats were necessary, so a new and moduled design of UDA was created in summer 2013. The UDA's modular architecture is depicted in Figure A.1. The main data storage is based on Pandas [96] but it is much extended. Pandas itself contributes a great deal to the time series analysis and statistical analysis of big datasets. Pandas, however, does not offer a sophisticated way of loading data from various sources. We have, therefore, designed a set of loaders customised for several sources that produce data in the Pandas DataFrame format. The list of currently supported loaders is depicted in Figure A.1.

The loaders were chosen based on needs of related projects in order to speed up all the necessary analysis. Besides the in time analysis, the UDA can combine data from installed inputs (loaders) and save them for further analysis in various formats in the form of datasets. These datasets can be saved for example in HDF5 (Hierarchical Data Format), CSV (Comma-Separated Values) and Excel formats. The UDA can also produce visualisations of the recorded data. The visualisation format is very similar to that of Matlab; however, the UDA uses matplotlib [95]. The visualisations can be saved in several formats, e.g. JPEG, PNG, SVG, EPS or PDF. The UDA supports two methods of data analysis – interactive and batch. In the interactive mode, IPython [97] is used, enabling us to provide examples easily or even conduct seminars on how to work with the UDA. The same commands can be saved in a file and called just once in the batch mode. The batch mode is generally useful for performing repetitive data analyses on several datasets. The interactive mode is more useful for prototyping.

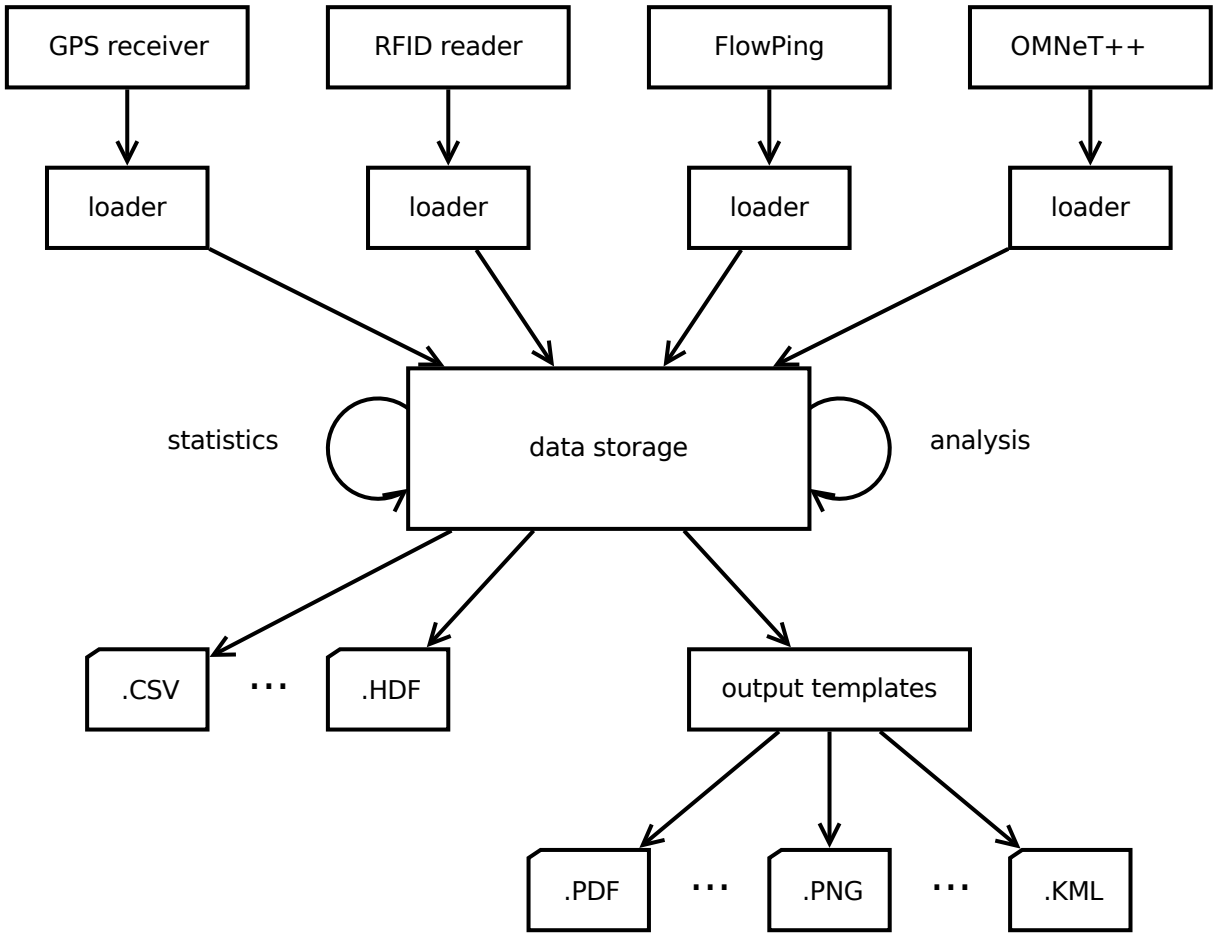


Figure A.1: UDA system - base topology

A.1 Requirements

The UDA can be operated on all platforms where Python can be run. The UDA is very light on the memory, as it was designed for embedded devices, however, the amount of analysed data is limited by the size of the memory installed on the embedded platform. Basic analysis can be performed on an embedded platform with the x86 or ARM CPU architecture. The required memory is 128 MB RAM. For very complex analyses, a considerable amount of memory and multicore CPUs must be used.

A.2 Common use

Some information about the use in common network environment:

- Analysis of system logs.
- Visualization and analysis of measurement data from sensor networks.
- Visualization and analysis of network simulator results.

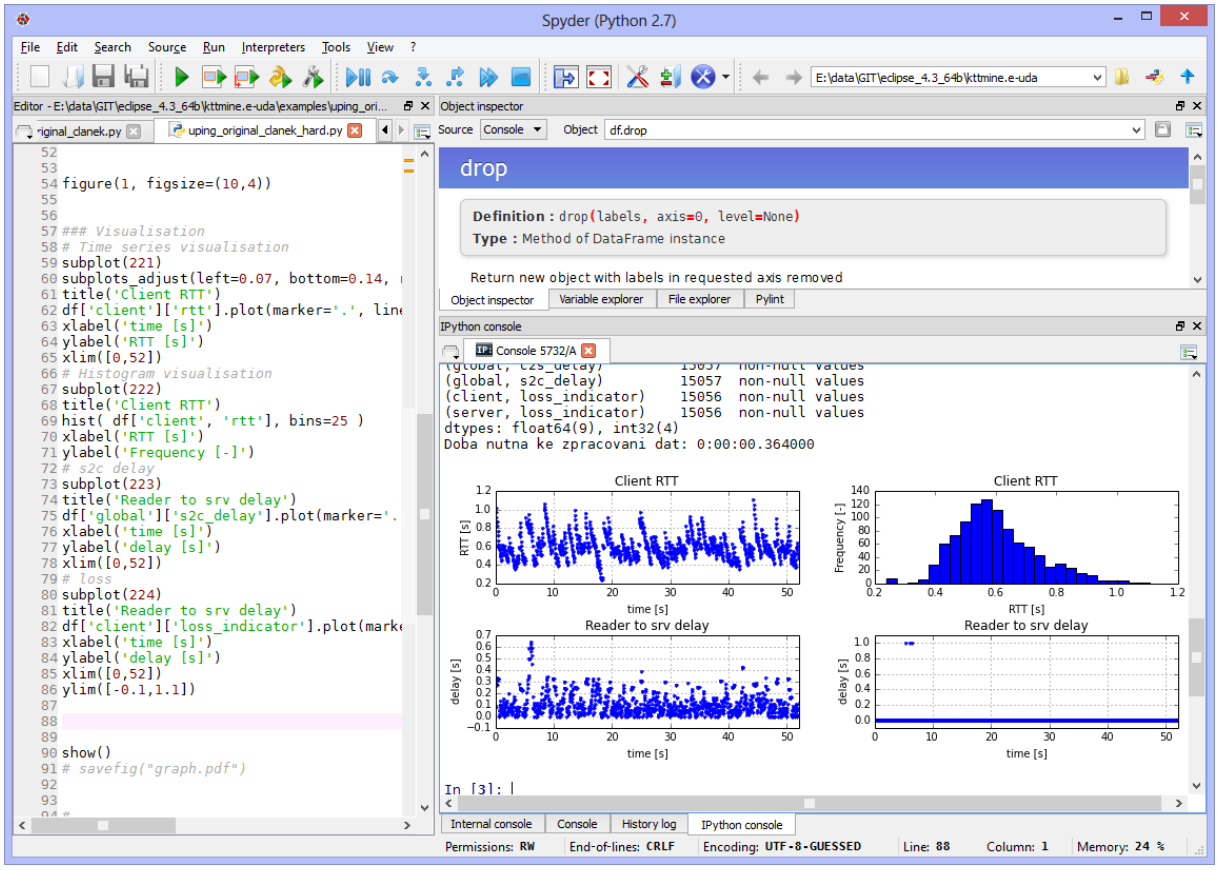


Figure A.2: UDA system - base topology

- Visualization and analysis of long-term measurements etc.

A.3 FlowPing use

The previous section, see Section A.2, tackles a common workflow. This section contains examples. The example illustrates a batch script:

```
from uda.loaders.flowping import FlowPing
from pylab import *

# Input data
fs='flowping-csv-sample/sample-server.csv'
fc='flowping-csv-sample/sample-client.csv'

# Data loading
fp= FlowPing(server_log=fs, client_log=fc)
df= fp.getDataFrame()
```

```
### Visualisation
# Time serie visualisation
subplot(121)
title('Client RTT')
df['client']['RTT'].plot()

# Histogram visualisation
subplot(122)
title('Client RTT')
hist( df['client', 'RTT'], bins=25 )

# Saving data in Excel format
df.to_excel("data.xlsx")

# Saving the visualisation
savefig("graph.pdf")
```

This script show how to read the data from FlowPing, and how to obtain data in the DataFrame format – *df* and how to visualize the data. The visualisation will be straightforward for Matlab expert users. The last line of the example script saves the visualisation into PDF. The visualisation (with some minor tweaks) is captured in Figure A.3.

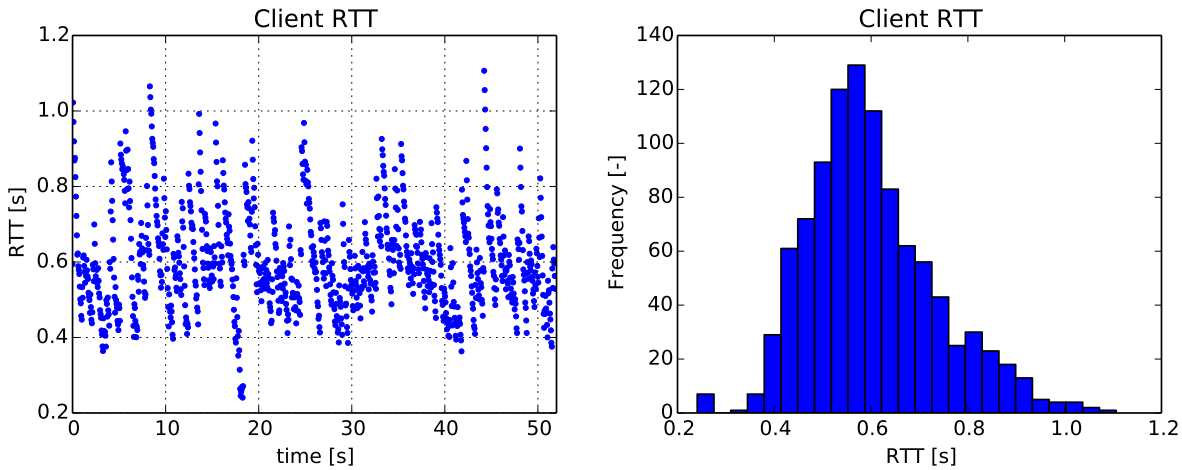


Figure A.3: Example of FlowPing visualisation from the UDA – *graph.pdf*.

The UDA was used with OMNeT++ simulations on Computer clusters. The analysis using UDA was carried out right after the simulation, so big amount of raw data generated by OMNeT++ was not necessary to save. Only the information in data were saved for further analysis.

A.4 Licensing

The UDA framework is developed as an open source project covered by the General Public License (GPL).

Bibliography

[1] ITU-T, "10-Gigabit-capable passive optical networks (XG-PON)", International Telecommunication Union, Geneva, Recommendation G.987.1, Nov. 2010.

[2] Cisco, *Cisco visual networking index: forecast and methodology, 2013–2018*, Last visit on Sep. 2014, 2014. [Online]. Available: http://bit.ly/P1eXkv.

[3] S. Bandyopadhyay, *Dissemination of Information in Optical Networks:: From Technology to Algorithms*, ser. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2007, ISBN: 9783540728757. [Online]. Available: http://books.google.cz/books?id=XaBzyH165GsC.

[4] C. Kallo, V. Lopez, J. Dunne, O. Gonzalez De Dios, M. Basham, and J. Fernandez-Palacios, "Benefits of optical packet switching for router by-pass in metro networks", in *Future Network & Mobile Summit (FutureNetw), 2012*, 2012, pp. 1–8. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6294235.

[5] A. A. M. Saleh and J. M. Simmons, "Technology and architecture to enable the explosive growth of the internet", *Communications Magazine, IEEE*, vol. 49, no. 1, pp. 126–132, 2011, ISSN: 0163-6804. DOI: 10.1109/MCOM.2011.5681026.

[6] W. Vereecken, W. Van Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester, "Power consumption in telecommunication networks: overview and reduction strategies", vol. 49, no. 6, pp. 62–69, 2011. DOI: 10.1109/MCOM.2011.5783986. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5783986.

[7] W. Vereecken, W. Van Heddeghem, B. Puype, D. Colle, M. Pickavet, and P. Demeester, "Optical networks: how much power do they consume and how can we optimize this?", in *Optical Communication (ECOC), 2010 36th European Conference and Exhibition on*, 2010, pp. 1–4. DOI: 10.1109/ECOC.2010.5621575. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5621575.

[8] A. Jourdan, F. Bakhti, L. Berthelon, F. Bruyere, M. Chbat, D. Chiaroni, C. Drion, G. Eilenberger, M. Garnot, F. Masetti, P. Perrier, and M. Renaud, "Key building blocks for high-capacity WDM photonic transport networks", vol. 16, no. 7, pp. 1286–1297, 1998. DOI: 10.1109/49.725196. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=725196.

[9] D. Monoyios and K. Vlachos, "Emulating lossless, one-way signaling protocols in OBS networks with traffic prediction", in *Optical Network Design and Modeling, 2008. ONDM 2008. International Conference on*, 2008, pp. 1–6. DOI: 10.1109/ONDM.2008.4578404. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4578404.

[10] W. Van Heddeghem, M. Parker, S. Lambert, W. Vereecken, B. Lannoo, D. Colle, M. Pickavet, and P. Demeester, "Using an analytical power model to survey power saving approaches in backbone networks", in *Networks and Optical Communications (NOC), 2012 17th European Conference on*, 2012, pp. 1–6. DOI: 10.1109/NOC.2012.6249942. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6249942.

[11] J. Baliga, R. Ayre, K. Hinton, W. Sorin, and R. Tucker, "Energy consumption in optical IP networks", vol. 27, no. 13, pp. 2391–2403, 2009. DOI: 10.1109/JLT.2008.2010142. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4815495.

[12] J. Simmons, "Network design in realistic "all-optical" backbone networks", *Communications Magazine, IEEE*, vol. 44, no. 11, pp. 88 –94, 2006, Podložení nutnosti regenerátorů, ISSN: 0163-6804. DOI: 10.1109/MCOM.2006.248170.

[13] C. Qiao and M. Yoo, "Optical burst switching (OBS) - a new paradigm for an optical internet", *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69–84, 1999.

[14] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: an approach to high bandwidth optical WAN's", *Communications, IEEE Transactions on*, vol. 40, no. 7, pp. 1171–1182, 1992, ISSN: 0090-6778. DOI: 10.1109/26.153361.

[15] T. El-Bawab, *Optical Switching*. Springer, 2008, ISBN: 9780387291598. [Online]. Available: http://www.google.cz/books?id=6f4c_IfbElMC.

[16] M. H. Phung, K. C. Chua, G. Mohan, M. Motani, and T. C. Wong, "An absolute QoS framework for loss guarantees in Optical Burst-Switched Networks", *Communications, IEEE Transactions on*, vol. 55, no. 6, pp. 1191–1201, 2007, ISSN: 0090-6778. DOI: 10.1109/TCOMM.2007.898846.

[17] M. Klinkowski, "Offset time-emulated architecture for optical burst switching - modelling and performance evaluation", PhD thesis, 2007.

[18] D. Hunter, M. Nizam, M. Chia, I. Andonovic, K. Guild, A. Tzanakaki, M. O'Mahony, L. Bainbridge, M. Stephens, R. Penty, and I. White, "WASPNET: a wavelength switched packet network", *Communications Magazine, IEEE*, vol. 37, no. 3, pp. 120–129, 1999, ISSN: 0163-6804. DOI: 10.1109/35.751509.

[19] P. Chu, C. Lee, and S. Park, "MEMS: the path to large optical crossconnects", *Communications Magazine, IEEE*, vol. 40, no. 3, pp. 80–87, 2002, ISSN: 0163-6804. DOI: 10.1109/35.989762.

[20] J. Cheyns, C. Develder, E. Van Breusegem, A. Ackaert, M. Pickavet, and P. Demeester, "Routing in an awg-based optical packet switch", English, *Photonic Network Communications*, vol. 5, no. 1, pp. 69–80, 2003, ISSN: 1387-974X. DOI: 10.1023/A:1021005913665. [Online]. Available: http://dx.doi.org/10.1023/A%3A1021005913665.

[21] R. S. Tucker and W. D. Zhong, *Photonic packet switching: an overview*, 1999.

[22] K. Nashimoto, N. Tanaka, M. LaBuda, D. Ritums, J. Dawley, M. Raj, D. Kudzuma, and T. Vo, "High-speed plzt optical switches for burst and packet switching", in *Broadband Networks, 2005. BroadNets 2005. 2nd International Conference on*, 2005, 1118–1123 Vol. 2. DOI: 10.1109/ICBN.2005.1589732.

[23] G. Shen and R. S. Tucker, "Translucent optical networks: the way forward [topics in optical communications]", *Communications Magazine, IEEE*, vol. 45, no. 2, pp. 48 –54, 2007, ISSN: 0163-6804. DOI: 10.1109/MCOM.2007.313394.

[24] P. Pedroso, J. Sole-Pareta, D. Careglio, and M. Klinkowski, "Integrating GMPLS in the OBS networks control plane", in *Transparent Optical Networks, 2007. ICTON '07. 9th International Conference on*, vol. 3, 2007, pp. 1–7. DOI: 10.1109/ICTON.2007.4296231.

[25] B. Mukherjee, *Optical WDM networks*. Springer Science, 2006.

[26] M. Klinkowski, J. Pedro, D. Careglio, M. Pióro, J. Pires, P. Monteiro, and J. Solé-Pareta, "An overview of routing methods in optical burst switching networks", *Optical Switching and Networking*, vol. 7, no. 2, pp. 41 –53, 2010, ISSN: 1573-4277. DOI: 10.1016/j.osn.2010.01.001.

[27] B. Jaumard, "Operation research tools and methodology for the design and provisioning of survivable optical networks", in *Optical Network Design and Modeling (ONDM), 2013 17th International Conference on*, 2013, pp. 53–58.

[28] M. D. Leenheer and *et al.*, "A view on enabling-consumer oriented grids through optical burst switching", *IEEE Communications Magazine*, pp. 124–131, 2006, ISSN: 0163-6804. DOI: 10.1109/MCOM.2006.1607875.

[29] Intune, *Optical packet switched transport*, http://www.intunenetworks.com, Last visit on Jan. 2014, 2012.

[30] Y. Liu, H. Guo, T. Tsuritani, Y. Yin, J. Wu, X. Hong, J. Lin, and M. Suzuki, "Dynamic provisioning of self-organized consumer grid services over integrated OBS/WSON networks", *Journal of Lightwave Technology*, vol. 30, no. 5, pp. 734–753, 2012, ISSN: 0733-8724. DOI: 10.1109/JLT.2011.2180508.

[31] *Organizational Psychology: An experimental approach. Hauptbd.* Prentice-Hall, 1979. [Online]. Available: http://books.google.cz/books?id=CbCCnAEACAAJ.

[32] V. Chvatal, *Linear Programming*. Freeman, 1983.

[33] J. Daigle, *Queueing Theory with applications to packet telecommunication*. Springer Science + Business Media, 2005.

[34] *OMNeT++ user manual*, 2014. [Online]. Available: http://www.omnetpp.org/doc/omnetpp/manual/usman.html#sec101.

[35] M. Bazaraa, J. Jarvis, and H. Sherali, *Linear Programming and Network Flows*. Wiley, 2011, ISBN: 9780471703761. [Online]. Available: http://books.google.ca/books?id=FykSXKGEeZQC.

[36] L. R. Ford and D. R. Fulkerson, "Maximal flow through a network", *CANADIAN JOURNAL OF MATHEMATICS*, pp. 399–404, 1956.

[37] Cplex, *IBM ILOG CPLEX 12.0 Optimization Studio*, IBM, 2011.

[38] T. Battestilli and H. Perros, "An introduction to optical burst switching", *Communications Magazine, IEEE*, vol. 41, no. 8, S10–S15, 2003, ISSN: 0163-6804. DOI: 10.1109/MCOM.2003.1222715.

[39] T. Battestilli and H. Perros, "Optical burst switching for the next generation Internet", *Potentials, IEEE*, vol. 23, no. 5, pp. 40–43, 2005, ISSN: 0278-6648. DOI: 10.1109/MP.2005.1368915.

[40] T. Venkatesh and C. Murthy, *An Analytical Approach to Optical Burst Switched Networks*. Springer, 2010.

[41] J. Jue and V. Vokkarane, *Optical Burst Switched Networks*, ser. Optical Networks. Springer, 2006, ISBN: 9780387237602. [Online]. Available: https://books.google.cz/books?id=wFHHgkmHNu0C.

[42] D. Chiaroni, "Optical packet add/drop multiplexers for packet ring networks", in *Optical Communication, 2008. ECOC 2008. 34th European Conference on*, 2008, pp. 1–4. DOI: 10.1109/ECOC.2008.4729412.

[43] D. Chiaroni, C. Simonneau, M. Salsi, G. Buforn, S. Etienne, H. Mardoyan, J. Simsarian, and J.-C. Antona, "Optical packet ring network offering bit rate and modulation formats transparency", in *Optical Fiber Communication (OFC), collocated National Fiber Optic Engineers Conference, 2010 Conference on (OFC/NFOEC)*, 2010, pp. 1–3.

[44] M. Phung, K. Chua, G. Mohan, M. Motani, and T. Wong, "The streamline effect in OBS networks and its application in load balancing", in *International Conference on Broadband Networks - BROADNETS*, vol. 1, 2005, pp. 283–290. DOI: 10.1109/ICBN.2005.1589625.

[45] S. Cao, N. Deng, T. Ma, J. Qi, X. Shi, J. He, and J. Zhou, "An optical burst ring network featuring sub-wavelength- and wavelength-granularity grooming", in *Photonics Global Conference (PGC), 2010*, 2010, pp. 1–3. DOI: 10.1109/PGC.2010.5706067.

[46] M. Duser and P. Bayvel, "Analysis of a dynamically wavelength-routed optical burst switched network architecture", *Lightwave Technology, Journal of*, vol. 20, no. 4, pp. 574–585, 2002, ISSN: 0733-8724. DOI: 10.1109/50.996576.

[47] J. Dunne, T. Farrell, and J. Shields, "Optical packet switch and transport: a new metro platform to reduce costs and power by 50deterministic performance levels", in *Transparent Optical Networks, 2009. ICTON '09. 11th International Conference on*, 2009, pp. 1–5. DOI: 10.1109/ICTON.2009.5185042.

[48] T. Coutelen, "Loss-free architectures in optical burst switched networks for a reliable and dynamic optical layer", PhD thesis, Concordia University, 2010.

[49] T. Coutelen, B. Jaumard, and G. Hébuterne, "An enhanced train assembly policy for lossless OBS with CAROBS", in *Communication Networks and Services Research Conference (CNSR)*, 2010, pp. 61 –68. DOI: 10.1109/CNSR.2010.21.

[50] F. Pavon-Marino and F. Neri, "On the myths of optical burst switching", *IEEE Transactions on Communications*, vol. 59, no. 9, pp. 2574–2584, 2011, ISSN: 0090-6778. DOI: 10.1109/TCOMM.2011.063011.100192.

[51] C. Raffaelli and M. Savi, "Hybrid contention resolution in optical switching fabric with QoS traffic", in *International Conference on Broadband Networks - BROADNETS*, 2009, pp. 1–8.

[52] O. Pedrola, D. Careglio, M. Klinkowski, and J. Sole-Pareta, "Regenerator placement strategies for translucent OBS networks", *Journal of Lightwave Technology*, vol. 29, no. 22, pp. 3408 –3420, 2011, ISSN: 0733-8724. DOI: 10.1109/JLT.2011.2168806.

[53] M. Kozak, B. Jaumard, and L. Bohac, "On regenerator placement in loss-less optical burst switching networks", in *36th International Conference on Telecommunications and Signal Processing (TSP)*, 2013, pp. 311–315. DOI: 10.1109/TSP.2013.6613942.

[54] A. A. Triki, P. Gavignet, B. Arzur, E. Le Rouzic, and A. Gravey, "Bandwidth allocation schemes for a lossless optical burst switching", in *Conference on Optical Network Design and Modeling - ONDM*, 2013, pp. 205–210.

[55] B. Wang and N. Lella, "Dynamic contention resolution in optical burst switched networks with partial wavelength conversion and fiber delay lines", in *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, vol. 3, 2004, 1862–1866 Vol.3. DOI: 10.1109/GLOCOM.2004.1378312.

[56] M. Dutta and V. Chaubey, "Contention resolution in optical burst switching (OBS) network: a time domain approach", in *International Conference on Fiber Optics and Photonics (PHOTONICS)*, 2012, pp. 1–3.

[57] H. Vu and M. Zukerman, "Blocking probability for priority classes in optical burst switching networks", *IEEE Communications Letters*, vol. 6, no. 5, pp. 214–216, 2002, ISSN: 1089-7798. DOI: 10.1109/4234.1001668.

[58] P. Delesques, T. Bonald, G. Froc, P. Ciblat, and C. Ware, "Enhancement of an optical burst switch with shared electronic buffers", in *Conference on Optical Network Design and Modeling - ONDM*, 2013, pp. 137–142.

[59] L. L. Velasco, M. Klinkowski, M. Ruiz, and J. Comellas, "Modeling the routing and spectrum allocation problem for flexgrid optical networks", *Photonic Network Communications*, vol. 24, no. 3, pp. 177–186, 2012.

[60] B. Vignac, B. Jaumard, and F. Vanderbeck, "Hierarchical optimization procedure for traffic grooming in WDM optical networks", in *Optical Network Design and Modeling, 2009. ONDM 2009. International Conference on*, 2009, pp. 1–6.

[61] J.-Q. Hu and B. Leida, "Traffic grooming, routing, and wavelength assignment in optical WDM mesh networks", in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 1, 2004, pp. – 501. DOI: 10.1109/INFCOM.2004.1354521.

[62] Z. Le, M. Fu, and W. Dong, "Gradient projection based RWA algorithm for OBS network", in *7th International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP)*, 2010, pp. 272–277.

[63] J. Triay and C. Cervello-Pastor, "Topology analysis of auto load-balancing RWA in optical burst-switched networks", in *20th Annual Wireless and Optical Communications Conference (WOCC)*, 2011, pp. 1–6. DOI: 10.1109/WOCC.2011.5872289.

[64] T. Coutelen, G. Hebuterne, and B. Jaumard, "An OBS RWA formulation for asynchronous loss-less transfer in OBS networks", in *International Workshop on High Performance Switching and Routing - HPSR*, 2009, pp. 1–6. DOI: 10.1109/HPSR.2009.5307423.

[65] B. Jaumard, C. Meyer, and B. Thiongane, "Comparison of ILP formulations for the RWA problem", *Optical Switching and Networking*, vol. 4, no. 3-4, pp. 157–172, 2007.

[66] ——, "On column generation formulations for the RWA problem", *Discrete Applied Mathematics*, vol. 157, pp. 1291–1308, 6 2009.

[67] A. Betker, C. Gerlach, R. Hülsermann, M. Jäger, M. Barry, S. Bodamer, J. Späth, C. Gauger, and K. M, "Reference transport network scenarios", *Technical report German Ministry of Education and Research within the MultiTeraNet*, p. 16, 2004.

[68] S. Wankhade and S. Kambale, "An evolutionary approach for LAUC scheduler in optical burst switching networks", *International Journal of Applied Information Systems*, vol. 2, no. 8, pp. 1–4, 2012.

[69] A. A. M. Saleh and J. M. Simmons, "All-Optical Networking – Evolution, Benefits, Challenges, and Future Vision", *Proceedings of the IEEE*, vol. 100, no. 5, pp. 1105–1117, 2012, ISSN: 0018-9219. DOI: 10.1109/JPROC.2011.2182589.

[70] A. Barradas and M. d. C. Medeiros, "Pre-planned optical burst switched routing strategies considering the streamline effect", English, *Photonic Network Communications*, vol. 19, no. 2, pp. 161–169, 2010, ISSN: 1387-974X. DOI: 10.1007/s11107-009-0221-y. [Online]. Available: http://dx.doi.org/10.1007/s11107-009-0221-y.

[71] A. Gumaste and I. Chlamtac, "Light-trails: a novel conceptual framework for conducting optical communications", in *High Performance Switching and Routing, 2003, HPSR. Workshop on*, 2003, pp. 251–256. DOI: 10.1109/HPSR.2003.1226714. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1226714.

[72] M. Kozak, B. Jaumard, and L. Bohac, "On the highly stable performance of loss-free optical burst switching networks", *Submitted for publication (Cahiers du GERAD # G-2014-27)*, 21 pages, 2014.

[73] C. Qiao, W. Wei, and X. Liu, "Extending generalized multiprotocol label switching (GMPLS) for polymorphous, agile, and transparent optical networks (PATON)", *IEEE Communications Letters*, vol. 44, pp. 104–114, 2006.

[74] A. Gumaste and I. Chlamtac, "Light-trails: an optical solution for IP transport", *Journal of Optical Networking*, vol. 3, no. 5, pp. 261–281, 2004.

[75] A. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition Techniques in Mathematical Programming*, ser. Engineering and Science Applications. Springer, 2006.

[76] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly, "SNDlib 1.0–Survivable Network Design Library", English, *Networks*, vol. 55, no. 3, pp. 276–286, 2010. DOI: 10.1002/net.20371. [Online]. Available: http://www3.interscience.wiley.com/journal/122653325/abstract.

[77] C. Chan, *Optical Performance Monitoring: Advanced Techniques for Next-Generation Photonic Networks*. Elsevier Science, 2010, ISBN: 9780080959177. [Online]. Available: http://books.google.ca/books?id=PvCN_biNky0C.

[78] M. Channegowda, R. Nejabati, and D. Simeonidou, "Software-defined optical networks technology and infrastructure: enabling software-defined optical network operations [invited]", *Optical Communications and Networking, IEEE/OSA Journal of*, vol. 5, no. 10, A274–A282, 2013, ISSN: 1943-0620. DOI: 10.1364/JOCN.5.00A274.

[79] P. Bhaumik, S. Zhang, P. Chowdhury, S.-S. Lee, J. Lee, and B. Mukherjee, "Software-defined optical networks (sdons): a survey", English, *Photonic Network Communications*, vol. 28, no. 1, pp. 4–18, 2014, ISSN: 1387-974X. DOI: 10.1007/s11107-014-0451-5. [Online]. Available: http://dx.doi.org/10.1007/s11107-014-0451-5.

[80] P. Newman, W. Edwards, R. Hinden, E. Hoffman, F. C. Liaw, T. Lyon, and G. Minshall, *Ipsilon's General Switch Management Protocol Specification Version 1.1*, RFC 1987 (Informational), Updated by RFC 2297, Internet Engineering Task Force, Aug. 1996. [Online]. Available: `http://www.ietf.org/rfc/rfc1987.txt`.

[81] S. Sezer, S. Scott-Hayward, P. Chouhan, B. Fraser, D. Lake, J. Finnegan, N. Viljoen, M. Miller, and N. Rao, "Are we ready for sdn? implementation challenges for software-defined networks", vol. 51, no. 7, pp. 36–43, 2013. DOI: `10.1109/MCOM.2013.6553676`. [Online]. Available: `http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6553676`.

[82] M. Casado, M. J. Freedman, J. Pettit, J. Luo, N. McKeown, and S. Shenker, "Ethane: taking control of the enterprise", *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 4, pp. 1–12, Aug. 2007, ISSN: 0146-4833. DOI: `10.1145/1282427.1282382`. [Online]. Available: `http://doi.acm.org/10.1145/1282427.1282382`.

[83] Open Networking Foundation, "Openflow switch specification 1.5.0", Tech. Rep., 2014. [Online]. Available: `https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.5.0.noipr.pdf`.

[84] A. Patel, P. Ji, and T. Wang, "Qos-aware optical burst switching in openflow based software-defined optical networks", in *Optical Network Design and Modeling (ONDM), 2013 17th International Conference on*, 2013, pp. 275–280.

[85] O. N. Foundation, "Software-defined networking: the new norm for networks", Tech. Rep., 2012. [Online]. Available: `https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/wp-sdn-newnorm.pdf`.

[86] H. Zang and J. P. Jue, "A review of routing and wavelength assignment approaches for wavelength-routed optical wdm networks", *Optical Networks Magazine*, vol. 1, pp. 47–60, 2000.

[87] C. Metz, "IP anycast point-to-(any) point communication", vol. 6, no. 2, pp. 94–98, 2002. DOI: `10.1109/4236.991450`. [Online]. Available: `http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=991450`.

[88] *Ieee standard for local and metropolitan area networks - virtual bridged local area networks- amendment 2: vlan classification by protocol and port*, IEEE Std 802.1v-2001, IEEE, 2001.

[89] Open Networking Foundation, "Openflow switch specification 1.4.0", Tech. Rep., 2013. [Online]. Available: `https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.4.0.pdf`.

[90] T. Hégr, L. Boháč, Z. Kocur, M. Vozňák, and P. Chlumský, "Methodology of the direct measurement of the switching latency", English, *Przeglad Elektrotechniczny*, vol. 89, no. 7/2013, pp. 59–63, 2013, ISSN: 0033-2097. [Online]. Available: `http://pe.org.pl/articles/2013/7/13.pdf`.

[91] L. Liu, D. Zhang, T. Tsuritani, R. Vilalta, R. Casellas, L. Hong, I. Morita, H. Guo, J. Wu, R. Martinez, and R. Munoz, "Field trial of an openflow-based unified control plane for multilayer multigranularity optical switching networks", *Lightwave Technology, Journal of*, vol. 31, no. 4, pp. 506–514, 2013, ISSN: 0733-8724. DOI: `10.1109/JLT.2012.2212179`.

[92] S. Huang, K. Kitayama, F. Cugini, F. Paolucci, A. Giorgetti, L. Valcarenghi, and P. Castoldi, "An experimental analysis on OSPF-TE convergence time", *Proc. SPIE*, vol. 7137, pp. 713 728–713728–11, 2008. DOI: 10.1117/12.803224. [Online]. Available: http://dx.doi.org/10.1117/12.803224.

[93] G Shen, S. Bose, T. Cheng, C Lu, and T. Chai, "Efficient heuristic algorithms for light-path routing and wavelength assignment in WDM networks under dynamically varying loads", *Computer Communications*, vol. 24, no. 3 - 4, pp. 364 –373, 2001, ISSN: 0140-3664. DOI: http://dx.doi.org/10.1016/S0140-3664(00)00236-X. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S014036640000236X.

[94] AT&T, *At&t vision alignment challenge technology survey*, 2013. [Online]. Available: http://soc.att.com/11iSBRf.

[95] J. D. Hunter, "Matplotlib: a 2d graphics environment", *Computing In Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[96] Python Data Analysis Library, *Open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language.* http://pandas.pydata.org/ (accessed on 2014-08-05), 2014.

[97] F. Pérez and B. E. Granger, "IPython: a system for interactive scientific computing", *Computing in Science and Engineering*, vol. 9, no. 3, pp. 21–29, May 2007, ISSN: 1521-9615. DOI: 10.1109/MCSE.2007.53. [Online]. Available: http://ipython.org.